Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
	Approved for public release;
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	distribution unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
TR-88-0007	

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Frank J. Seiler Research Lab	FJSRL/NC	

6c. ADDRESS (City, State and ZIP Code)	7b. ADDRESS (City, State and ZIP Code)
USAF Academy	
Colorado Springs CO 80840-6528	

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AF Office of Scientific Research	AFOSR	

8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS			
Bldg 410	PROGRAM ELEMENT NO.	PROJECT NO	TASK NO	WORK UNIT NO
Bolling AFB DC 20332				

11. TITLE (Include Security Classification) MOPAC Manual Fifth Edition, A General Molecular Orbital Package	61102F	2303	F3	07

12. PERSONAL AUTHOR(S)
James J.P. Stewart

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Tech Report	FROM Oct 87 TO Nov 88	88 Nov	176

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	MOPAC, MNDO, AM1, PM3, Semi-empirical Calculations
0704	1901	2102	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

MOPAC is a general-purpose, semi-empirical molecular orbital program for the study of chemical reactions involving molecules, ions, and linear polymers. It implements the semi-empirical Hamiltonians MNDO, AM1, PM3, and MINDO/3 and combines the calculations of vibrational spectra, thermodynamic quantities, isotopic substitution effects, time-dependent effects, and force constants in a fully integrated program. Elements parametrized at the MNDO level include H, Li, Be, B, C, N, O, F, Al, Si, P, S, Cl, Cr, Ge, Sn, Hg, Pb, and I. Within the electronic part of the calculation, molecular and localized orbitals excited states up to sextets, chemical bond indices, charges, etc. are computed. Both intrinsic and dynamic reaction coordinates can be calculated. A transition-state location routine and two transition-state optimizing routines are available for studying chemical reactions.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS ☐	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
James J.P. Stewart	(303)472-2655	FJSRL/NC

DD FORM 1473, 83 APR   EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

MOPAC   MANUAL

(Fifth Edition)

A GENERAL MOLECULAR ORBITAL PACKAGE

DTIC
COPY
INSPECTED
1

Written by
James J. P. Stewart,
Frank J. Seiler Research Laboratory
United States Air Force Academy
Colorado Springs, CO 80840

1

CONTENTS

CHAPTER 7   PROGRAM

CHAPTER 8   ERROR MESSAGES PRODUCED BY MOPAC

CHAPTER 9   CRITERIA

CHAPTER 10   DEBUGGING

CHAPTER 11   INSTALLING MOPAC

APPENDIX A   FORTRAN FILES

APPENDIX B   SUBROUTINE CALLS IN MOPAC

APPENDIX C   DESCRIPTION OF SUBROUTINES IN MOPAC

APPENDIX D   HEATS OF FORMATION OF SOME MNDO, PM3 AND AM1 COMPOUNDS

APPENDIX E   REFERENCES

1 FORWARD BY PROF. MICHAEL J. S. DEWAR

"MOPAC is the present culmination of a continuing project that started twenty years ago, directed to the development of quantum mechanical procedures simple enough, and accurate enough, to be useful to chemists as an aid in their own research. A historical account of this development, with references, has appeared [1]. The first really effective treatment was MINDO/3 [2], which is still useful in various areas of hydrocarbon chemistry but ran into problems with heteroatoms. This was succeeded by MNDO [3] and more recently by AM1 [4] which seems to have overcome most of the deficiencies of its predecessors at no cost in computing time.

Our computer programs steadily evolved with the development of new algorithms. In addition to the basic programs for the SCF calculations and geometry optimization, programs were developed for calculating vibration frequencies [5], thermodynamic parameters [6], kinetic isotope effects [7], linear polymers [8], polarizabilities and hyperpolarizabilities [9,10], and SCF-CI calculations [11]. While this disjointed collection of programs served its purpose, it was inconvenient and time consuming to use. A major step was the integration [12] of most of these into a single unified program [MOPAC] with a greatly simplified input. The individual programs were also rewritten in a more efficient form so that the computing time reported for most calculations has now been halved. In its present form MOPAC is impressively easy to use and it contains options for nearly all the applications where our procedures have been found useful."

Michael J.S. Dewar, January 1987

REFERENCES

(1) Dewar, M.J.S., J. Mol. Struct., 100, 41 (1983).

(2) Dewar, M.J.S.; Bingham, R.C.; Lo, D.H., J. Am. Chem. Soc., 97, 1285 (1975).

(3) Dewar, M.J.S.; Thiel, W., J. Am. Chem. Soc., 99, 4899 (1977).

(4) Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.J.P., J. Am. Chem. Soc.; 107, 3902 (1985).

(5) Dewar, M.J.S.; Ford, G.P., J. Am. Chem. Soc., 99, 1685 (1977); Dewar, M.J.S.; Ford, G.; McKee, M.; Rzepa, H.S.; Yamaguchi, Y., J. Mol. Struct., 43, 135 (1978).

(6) Dewar, M.J.S.; Ford, G., J. Am. Chem. Soc., 99, 7822 (1977).

(7) Dewar, M.J.S.; Brown, S.B.; Ford, G.P.; Nelson, D.J.; Rzepa, H.S., J. Am. Chem. Soc., 100, 7832 (1978).

(8) Dewar, M.J.S.; Yamaguchi, Y.; Suck, S.H., Chem. Phys. Lett., 50,175,279 (1977).

(9) Dewar, M.J.S.; Bergman, J.G.; Suck, S.H.; Weiner, P.K., Chem. Phys. Lett., 38,226,(1976); Dewar, M.J.S.; Yamaguchi, Y.; Suck, S.H., Chem. Phys. Lett., 59, 541 (1978).

(10) Dewar, M.J.S.; Stewart, J.J.P., Chem. Phys. Lett., 111,416 (1984).

(11) Dewar, M.J.S.; Doubleday, C., J. Am. Chem. Soc., 100, 4935 (1978).

(12) Stewart, J.J.P. QCPE # 455.

2 UPDATES FROM VERSION 4.00

MOPAC is updated once a year. This is the best compromise between staying current and asking users to continuously change their software. Updates may be obtained from QCPE at the same cost as the original, or from sites that have a current copy. All VAX versions of MOPAC have the same QCPE number - 455; they are distinguished by version numbers. Users are recommended to update their programs at least once every two years, and preferably every year.

New Features of Version 5.0

1. A new method, MNDO-PM3, has been added. This method allows hypervalent systems to be calculated with accuracies comparable with non-hypervalent compounds.

2. In the DRC the option has been provided for the user to supply initial velocity vectors from the data-file. See keyword VELOCITY.

3. Time can now be specified as seconds (the default), minutes, hours, or days. See keyword T=.

4. The energy partition output has been rewritten so that each type of interaction is printed in matrix form. The code for this was written by Prof Tsuneo Hirano.

5. Hyperpolarizabilities can be calculated by invoking the keyword POLAR. The POLAR routine was completely rewritten by Dr. Henry Kurtz of Memphis State University.

6. The SHIFT option is now used by default. To prevent it being used, add the keyword SHIFT=0.0. The SHIFT option has been extensively re-written to allow the amount of damping of the SCF iterations to be determined by the rate of convergence. The basic methodology was described by A. V. Mitin -- see references. The effect of this is to allow systems that previously failed to go SCF to now converge.

7. Notwithstanding the improved SHIFT option, some systems still persist in failing to go SCF. If this is in a non-FORCE calculation, then if an SCF is almost achieved a warning will be printed and the calculation will continue. In a FORCE calculation the SCF must be completed to the precision selected, otherwise the second derivatives will not be sufficiently accurate.

8. The keyword GNORM now applies to both gradient minimizations as well as to the geometry optimization.

9. The meaning of LET has been generalized to now mean override default safety features.

10. Boron has been added to the AM1 set of elements: see references.

11. Zinc has been added to the MNDO set of elements: see references.

12. Some more warning messages are produced. In MECI, for example, if you specify the second triplet when a C.I.=2 is used, you will be told that the allowed states are three singlets and one triplet.

13. REPP and ROTATE have been completely re-written by Prof. Ernest R. Davidson of Indiana University. The new routines are easier to read and faster on a vector machine.

14. In a THERMO calculation, the heat of formation of the system relative to its elements in their standard state at 298K is printed.

ERRORS CORRECTED IN VERSION 5.0

1. Sometimes the geometry optimization routine would call for a sudden large change in the geometry. This would result in a nonsensical geometry, and the failure to achieve an SCF. All steps are now conservative.

2. In COMPFG some coordinates are printed if DEBUG and COMPFG are used. When the calculation was run in cartesian coordinates the output was incorrectly multiplied by 180/pi. This has been corrected.

3. Even when analytic derivatives were not used, channel 2 was touched (rewound). This gave rise to a spurious file FOR002.DAT. This file is now not created unless ANALYT is specified.

4. Systems of zero electrons are now allowed.

5. In a SADDLE calculation, if the number of atoms in the reactants is different from that in the products, the calculation is stopped.

6. In forming the inverse square root of the overlap matrix in the Mulliken population, a crash sometimes occurred due to the existence of negative molecular overlap integrals -- theoretically they cannot occur, but sometimes result due to round-off. This has been corrected.

7. All variables should now be initialized before use. The lack of initialization caused difficulties with certain computers.

8. In certain instances the bugfix to DIAG made in Version 4.0 slowed down the rate at which an SCF was achieved. This has been corrected by going back to the original formulation, but limiting the cosine to the range 1.0 to -1.0.

9. Under certain circumstances the "THREE ATOMS IN A STRAIGHT LINE" error message would be generated spuriously, and the calculation stopped. This has been corrected.

10. When PULAY is used in a FORCE calculation of an excited state, the SCF would sometimes lock onto the excited state. This was a result of FORCE needing the excited density matrix for the I.R. transition dipole calculation. Later on, when ITER is called, the density matrix is that of the excited state, and as PULAY is similar to the gradient minimization in that it converges on the nearest stationary point, the SCF that results is an excited state. This really messes up the FORCE calculation. To correct this, the old density matrix is reformed in ITER before the SCF is done.

Help with MOPAC

Telephone and mail support is given by the
Frank J. Seiler Research Laboratory on a time
permitting basis. If you need help, call
the Seiler MOPAC Consultant at

    (719) 472-2655

Similarly, mail should be addressed to

    MOPAC Consultant
    FJSRL/NC
    U.S. Air Force Academy CO 80840-6528
---

3 ACKNOWLEDGEMENTS

Acknowledgements

CHAPTER 1

DESCRIPTION OF MOPAC

MOPAC is a general-purpose semi-empirical molecular orbital package for the study of chemical structures and reactions. The semi-empirical Hamiltonians MNDO, MINDO/3, AM1, and PM3 are used in the electronic part of the calculation to obtain molecular orbitals, the heat of formation and its derivative with respect to molecular geometry. Using these results MOPAC calculates the vibrational spectra, thermodynamic quantities, isotopic substitution effects and force constants for molecules, radicals, ions, and polymers. For studying chemical reactions, a transition-state location routine and two transition state optimizing routines are available. For users to get the most out of the program, they must understand how the program works, how to enter data, how to interpret the results, and what to do when things go wrong.

While MOPAC calls upon many concepts in quantum theory and thermodynamics and uses some fairly advanced mathematics, the user need not be familiar with these specialized topics. MOPAC is written with the non-theoretician in mind. The input data are kept as simple as possible so users can give their attention to the chemistry involved and not concern themselves with quantum and thermodynamic exotica.

The simplest description of how MOPAC works is that the user creates a data-file which describes a molecular system and specifies what kind of calculations and output are desired. The user then commands MOPAC to carry out the calculation using that data-file. Finally the user extracts the desired output on the system from the output files created by MOPAC.

NOTES (1) This is the "fifth edition". MOPAC has undergone a steady expansion since its first release, and users of the earlier editions are recommended to familiarize themselves with the changes which are described in this manual. If any errors are found, or if MOPAC does not perform as described, please contact Dr. James J. P. Stewart, Frank J. Seiler Research Laboratory, U.S. Air Force Academy, Colorado Springs, CO 80840-6528. (2) MOPAC runs successfully on normal CDC, Data General, Gould, and Digital computers, and also on the CDC 205 and CRAY-XMP "supercomputers". The CRAY version has been partly optimized to take advantage of the CRAY architecture. Several versions exist for microcomputers such as the IBM PC-AT and XT, Zenith, etc.

1.1 SUMMARY OF MOPAC CAPABILITIES

1. MNDO, MINDO/3, AM1, and PM3 Hamiltonians.

2. Restricted Hartree-Fock (RHF) and Unrestricted Hartree-Fock (UHF) methods.

3. Extensive Configuration Interaction

   1. 100 configurations

   2. Singlets, Doublets, Triplets, Quartets, Quintets, and Sextets

   3. Excited states

   4. Geometry optimizations, etc., on specified states

4. Single SCF calculation

5. Geometry optimization

6. Gradient minimization

7. Transition state location

8. Reaction path coordinate calculation

9. Force constant calculation

10. Normal coordinate analysis

11. Transition dipole calculation

12. Thermodynamic properties calculation

13. Localized orbitals

14. Covalent bond orders

15. Bond analysis into sigma and pi contributions

16. One dimensional polymer calculation

17. Dynamic Reaction Coordinate calculation

18. Intrinsic Reaction Coordinate calculation

1.2 COPYRIGHT STATUS OF MOPAC

At the request of the Air Force Academy Law Department the following notice has been placed in MOPAC.

Notice of Public Domain nature of MOPAC

I recommend that a user obtain a copy by either copying it from an existing site or ordering an 'official' copy from the Quantum Chemistry Program Exchange, (QCPE), Department of Chemistry, Indiana University, Bloomington, Indiana, 47405. The cost covers handling only. Contact the Editor, Richard Counts at (812) 335-4784 for further details.

1.3 RELATIONSHIP OF AMPAC AND MOPAC

In 1985 MOPAC 3.0 and AMPAC 1.0 were submitted to QCPE for distribution. At that time, AMPAC differed from MOPAC in that it had the AM1 algorithm. Additionally, changes in some MNDO parameters in AMPAC made AMPAC results incompatable with MOPAC Versions 1-3. Subsequent versions of MOPAC, in addition to being more highly debugged than Version 3.0, also had the AM1 method. Such versions were compatable with AMPAC and with versions 1-3 of MOPAC.

In order to avoid confusion, all versions of MOPAC after 3.0 include journal references so that the user knows unambiguously which parameter sets were used in any given job.

Since 1985 AMPAC and MOPAC have evolved along different lines. In MOPAC I have endeavoured to provide a highly robust program, one with only a few new features, but which is easily portable and which can be relied upon to give precise, if not very exciting, answers. At Austin, the functionality of AMPAC has been enhanced by the research work of Prof. Dewar's group. AMPAC thus has functionalities not present in MOPAC. In publications, users should cite not only the program name but also the version number.

Commercial concerns have optimized both MOPAC and AMPAC for use on supercomputers. The quality of optimization and the degree to which the parent algorithm has been preserved differs between MOPAC and AMPAC and also between some machine specific versions. Different users may prefer one program to the other, based on considerations such as speed. Some

modifications of AMPAC run faster than some modifications of MOPAC, and vice versa, but if these are modified versions of MOPAC 3.0 or AMPAC 1.0, they represent the programming prowess of the companies doing the conversion, and not any intrinsic difference between the two programs.

Testing of these large algorithms is difficult, and several times users have reported bugs in MOPAC or AMPAC which were introduced after they were supplied by QCPE.

Cooperative Development of MOPAC

MOPAC has developed, and hopefully will continue to develop, by the addition of contributed code. As a policy, any supplied code which is incorporated into MOPAC will be described in the next release of the Manual, and the author or supplier acknowledged. In the following release only journal references will be retained. The objective is to produce a good program. This is obviously not a one-person undertaking, if it was, then the product would be poor indeed. Instead, as we are in a time of rapid change in computational chemistry, a time characterized by a very free exchange of ideas and code, MOPAC has been evolving by accretion. The unstinting and generous donation of intellectual effort speaks highly of the donors, however, with the rapid commercialization of computational chemistry software in the past few years, it is unfortunate but it seems unlikely that this idyllic state will continue.

1.4 PROGRAMS RECOMMENDED FOR USE WITH MOPAC

MOPAC is the core program of a series of programs for the theoretical study of chemical phenomena. This version is the third in an on-going development, and efforts are being made to continue its further evolution. In order to make using MOPAC easier, four other programs have also been written. Users of MOPAC are recommended to use all four programs. Efforts will be made to continue the development of these programs.

DRAW

DRAW, written by Maj. Donn Storch, USAF, and available through QCPE, is a powerful editing program specifically written to interface with MOPAC. Among the various facilities it offers are:

1. The on-line editing and analysis of a data file, starting from scratch or from an existing data file, an archive file, or from a results file.

2. The option of continuous graphical representation of the system being studied. Several types of terminals are supported, including DIGITAL, TEKTRONIX, and TERAK terminals.

3. The drawing of electron density contour maps generated by DENSITY on graphical devices.

4. The drawing of solid-state band structures generated by MOSOL.

5. The sketching of molecular vibrations, generated by a normal coordinate analysis.

DENSITY

DENSITY, written by Dr. James J. P. Stewart, and available through QCPE, is an electron-density plotting program. It accepts data-files directly from MOPAC, and is intended to be used for the graphical representation of electron density distribution, individual M.O.'s, and difference maps.

MOHELP

MOHELP, also available through QCPE, is an on-line help facility, written by Maj. Donn Storch and Dr. James J. P. Stewart, to allow non-VAX users access to the VAX HELP libraries for MOPAC, DRAW, and DENSITY.

MOSOL

MOSOL (Distributed by QCPE) is a full solid-state MNDO program written by Dr. James J. P. Stewart. In comparison with MOPAC, MOSOL is extremely slow. As a result, while geometry optimization, force constants, and other functions can be carried out by MOSOL, these slow calculations are best done using the solid-state facility within MOPAC. MOSOL should only be used for generating band-structures and densities of states, a task that MOPAC cannot perform.

1.5 THE DATA-FILE

This section is aimed at the complete novice -- someone who knows nothing at all about the structure of a MOPAC data-file.

First of all, there are at most four possible types of data-files for MOPAC, but the simplest data-file is the most commonly used. Rather than define it, two examples are shown below. An explanation of the geometry definitions shown in the examples is given in the chapter "GEOMETRY SPECIFICATION".

1.5.1 Example Of Data For Ethylene

```
Line   1 :       UHF PULAY MINDO3 VECTORS DENSITY LOCAL T=300
Line   2 :         EXAMPLE OF DATA FOR MOPAC
Line   3 :          MINDO/3 UHF CLOSED-SHELL D2D ETHYLENE
Line   4a:    C
Line   4b:    C    1.400118  1
Line   4c:    H    1.098326  1   123.572063  1
Line   4d:    H    1.098326  1   123.572063  1   180.000000  0   2  1  3
Line   4e:    H    1.098326  1   123.572063  1    90.000000  0   1  2  3
Line   4f:    H    1.098326  1   123.572063  1   270.000000  0   1  2  3
Line   5 :
```

As can be seen, the first three lines are textual. The first line
consists of keywords (here seven keywords are shown). These control the
calculation. The next two lines are comments or titles. The user might
want to put the name of the molecule and why it is being run on these two
lines.

These three lines are obligatory. If no name or comment is wanted,
leave blank lines. If no keywords are specified, leave a blank line. A
common error is to have a blank line before the keyword line: this error
is quite tricky to find, so be careful not to have four lines before the
start of the geometric data (lines 4a-4f in the example).   Whatever is
decided, the three lines, blank or otherwise, are obligatory.

The next set of lines defines the geometry. In the example, the
numbers are all neatly lined up; this is not necessary, but does make it
easier when looking for errors in the data. The geometry is defined in
lines 4a to 4f; line 5 terminates both the geometry and the data-file.
Any additional data, for example symmetry data, would follow line 5.

Summarizing, then, the structure for a MOPAC data-file is:

```
Line   1:   Keywords. (See chapter 2 on definitions of keywords)
Line   2:   Title of the calculation,  e.g. the name of the molecule or ion.
Line   3:   Other information describing the calculation.
Lines  4:   Internal or cartesian coordinates (See chapter on
            specification of geometry)
Line   5:   Blank line to terminate the geometry definition.
```

Other layouts for data-files involve additions to the simple layout.
These additions occur at the end of the data-file, after line 5. The
three most common additions are:

(a) Symmetry data: This follows the geometric data, and is ended by
a blank line.

(b) Reaction path: After all geometry and symmetry data (if any)
are read in, points on the reaction coordinate are defined.

(c) Saddle data: A complete second geometry is input. The second
geometry follows the first geometry and symmetry data (if any).

1.5.2 Example Of Data For Polytetrahydrofuran

The following example illustrates the data file for a four hour
polytetrahydrofuran calculation.   As you can see the layout of the data
is almost the same as that for a molecule, the main difference is in the
presence of the translation vector atom "Tv".

```
Line 1 :T=4H
Line 2 :        POLY-TETRAHYDROFURAN (C4 H8 O)2
Line 3 :
Line 4a:   C     0.000000  0     0.000000  0     0.000000  0   0  0  0
Line 4b:   C     1.551261  1     0.000000  0     0.000000  0   1  0  0
Line 4c:   O     1.401861  1   108.919034  1     0.000000  0   2  1  0
Line 4d:   C     1.401958  1   119.302489  1  -179.392581  1   3  2  1
Line 4e:   C     1.551074  1   108.956238  1   179.014664  1   4  3  2
Line 4f:   C     1.541928  1   113.074843  1   179.724877  1   5  4  3
Line 4g:   C     1.551502  1   113.039652  1   179.525806  1   6  5  4
Line 4h:   O     1.402677  1   108.663575  1   179.855864  1   7  6  5
Line 4i:   C     1.402671  1   119.250433  1  -179.637345  1   8  7  6
Line 4j:   C     1.552020  1   108.665746  1  -179.161900  1   9  8  7
Line 4k:   XX    1.552507  1   112.659354  1  -178.914985  1  10  9  8
Line 4l:   XX    1.547723  1   113.375266  1  -179.924995  1  11 10  9
Line 4m:   H     1.114250  1    89.824605  1   126.911018  1   1  3  2
Line 4n:   H     1.114708  1    89.909148  1  -126.650667  1   1  3  2
Line 4o:   H     1.123297  1    93.602831  1   127.182594  1   2  4  3
Line 4p:   H     1.123640  1    93.853406  1  -126.320187  1   2  4  3
Line 4q:   H     1.123549  1    90.682924  1   126.763659  1   4  6  5
Line 4r:   H     1.123417  1    90.679889  1  -127.033695  1   4  6  5
Line 4s:   H     1.114352  1    90.239157  1   126.447043  1   5  7  6
Line 4t:   H     1.114462  1    89.842852  1  -127.140168  1   5  7  6
Line 4u:   H     1.114340  1    89.831790  1   126.653999  1   6  8  7
Line 4v:   H     1.114433  1    89.753913  1  -126.926618  1   6  8  7
Line 4w:   H     1.123126  1    93.644744  1   127.030541  1   7  9  8
Line 4x:   H     1.123225  1    93.880969  1  -126.380511  1   7  9  8
Line 4y:   H     1.123328  1    90.261019  1   127.815464  1   9 11 10
Line 4z:   H     1.123227  1    91.051403  1  -125.914234  1   9 11 10
Line 4A:   H     1.113970  1    90.374545  1   126.799259  1  10 12 11
Line 4B:   H     1.114347  1    90.255788  1  -126.709810  1  10 12 11
Line 4C:   Tv   12.299490  1     0.000000  0     0.000000  0   1 11 10
Line 5 :   0     0.000000  0     0.000000  0     0.000000  0   0  0  0
```

Polytetrahydrofuran has a repeat unit of (C4 H8 O)2; i.e., twice the
monomer unit. This is necessary in order to allow the lattice to repeat
after a translation through 12.3 Angstroms. See the section on Solid
State Capability for further details.

Note the two dummy atoms on lines 4k and 4l. These are useful, but
not essential, for defining the geometry. The atoms on lines 4y to 4B
use these dummy atoms, as does the translation vector on line 4C.   The
translation vector has only the length marked for optimization. The
reason for this is also explained in the Background chapter.

CHAPTER 2

KEYWORDS

2.1 SPECIFICATION OF KEYWORDS

All control data are entered in the form of keywords, which form the first line of a data-file. A description of what each keyword does is given in Section 2-3. The order in which keywords appear is not important although they must be separated by a space. Some keywords can be abbreviated, allowed abbreviations are noted in Section 2-3 (for example 1ELECTRON can be entered as 1ELECT). However the full keyword is preferred in order to more clearly document the calculation and to obviate the possibility that an abbreviated keyword might not be recognized. If there is insufficient space in the first line for all the keywords needed, then consider abbreviating the longer words. One type of keyword, those with an equal sign, such as, BAR=0.05, may not be abbreviated, and the full word needs to be supplied.

Most keywords which involve an equal sign, such as SCFCRT=1.D-12 can, at the user's discretion, be written with spaces before and after the equal sign. Thus all permutations of SCFCRT=1.D-12, such as SCFCRT =1.D-12, SCFCRT = 1.D-12, SCFCRT= 1.D-12, SCFCRT = 1.D-12, etc. are allowed. Exceptions to this are T=, T-PRIORITY=, H-PRIORITY=, X-PRIORITY=, IRC=, DRC= and TRANS=. ' T=' cannot be abbreviated to ' T ' as many keywords start or end with a 'T'; for the other keywords the associated abbreviated keywords have specific meanings.

If two keywords which are incompatible, like UHF and C.I., are supplied, or a keyword which is incompatible with the species supplied, for instance TRIPLET and a methyl radical, then error trapping will normally occur, and an error message will be printed. This usually takes an insignificant time, so data are quickly checked for obvious errors.

2.2 FULL LIST OF KEYWORDS USED IN MOPAC

```
0SCF       - READ IN DATA, THEN STOP
1ELECTRON- PRINT FINAL ONE-ELECTRON MATRIX
1SCF       - DO ONE SCF AND THEN STOP
ANALYT     - USE ANALYTICAL DERIVATIVES OF ENERGY W.R.T. GEOMETRY
AM1        - USE THE AM1 HAMILTONIAN
BAR=n.n    - REDUCE BAR LENGTH BY A MAXIMUM OF n.n
BIRADICAL- SYSTEM HAS TWO UNPAIRED ELECTRONS
BONDS      - PRINT FINAL BOND-ORDER MATRIX
C.I.       - A MULTI-ELECTRON CONFIGURATION INTERACTION SPECIFIED
CHARGE=n   - CHARGE ON SYSTEM = n (e.g. NH4 => CHARGE=1)
COMPFG     - PRINT HEAT OF FORMATION CALCULATED IN COMPFG
CYCLES     - PERFORM MAXIMUM NUMBER OF CYCLES IN NLLSQ
DCART      - PRINT DETAILS OF WORKING IN DCART
DEBUG      - DEBUG OPTION TURNED ON
DEBUGPULAY PRINT DETAILS OF WORKING IN PULAY
DENOUT     - DENSITY MATRIX OUTPUT (CHANNEL 10)
DENSITY    - PRINT FINAL DENSITY MATRIX
DEP        - GENERATE FORTRAN CODE FOR PARAMETERS FOR NEW ELEMENTS
DEPVAR=n   - TRANSLATION VECTOR IS A MULTIPLE OF BOND-LENGTH
DERIV      - PRINT PART OF WORKING IN DERIV
DFORCE     - FORCE CALCULATION SPECIFIED, ALSO PRINT FORCE MATRIX.
DFP        - USE DAVIDON-FLETCHER-POWELL METHOD TO OPTIMIZE GEOMETRIES
DOUBLET    - RHF DOUBLET STATE REQUIRED
DRC        - DYNAMIC REACTION COORDINATE CALCULATION
DUMP=n     - WRITE RESTART FILES EVERY n SECONDS
ECHO       - DATA ARE ECHOED BACK BEFORE CALCULATION STARTS
EIGS       - PRINT ALL EIGENVALUES IN ITER
ENPART     - PARTITION ENERGY INTO COMPONENTS
ESR        - CALCULATE RHF UNPAIRED SPIN DENSITY
EXCITED    - OPTIMIZE FIRST EXCITED SINGLET STATE
EXTERNAL   - READ MNDO OR AM1 PARAMETERS OFF DISK
FILL=n     - IN RHF OPEN AND CLOSED SHELL, FORCE M.O. n
             TO BE FILLED
FLEPO      - PRINT DETAILS OF GEOMETRY OPTIMIZATION
FMAT       - PRINT DETAILS OF WORKING IN FMAT
FOCK       - PRINT LAST FOCK MATRIX
FORCE      - FORCE CALCULATION SPECIFIED
FULSCF     - FULL SCF CALCN'S TO BE DONE IN SEARCHES, AND DERIVATIVES
             WHEN NON-VARIATIONALLY OPTIMIZED WAVEFUNCTIONS USED
GEO-OK     - OVERRIDE INTERATOMIC DISTANCE CHECK
GNORM=n.n- EXIT WHEN GRADIENT NORM DROPS BELOW n.n
GRADIENTS- PRINT ALL GRADIENTS
GRAPH      - GENERATE FILE FOR GRAPHICS
HCORE      - PRINT DETAILS OF WORKING IN HCORE
H-PRIO     - HEAT OF FORMATION TAKES PRIORITY IN DRC
IRC        - INTRINSIC REACTION COORDINATE CALCULATION
ISOTOPE    - FORCE MATRIX WRITTEN TO DISK (CHANNEL 9 )
ITER       - PRINT DETAILS OF WORKING IN ITER
ITRY=N     - SET LIMIT OF NUMBER OF SCF ITERATIONS TO N.
KINETIC    - EXCESS KINETIC ENERGY ADDED TO DRC CALCULATION
```

```
LINMIN   - PRINT DETAILS OF LINE MINIMIZATION
LARGE    - PRINT EXPANDED OUTPUT
LET      - OVERRIDE CERTAIN SAFETY CHECKS
LOCALIZE - PRINT LOCALIZED ORBITALS
MINDO/3  - USE THE MINDO/3 HAMILTONIAN
MECI     - PRINT DETAILS OF MECI CALCULATION
MICROS   - USE SPECIFIC MICROSTATES IN THE C.I.
MMOK     - USE MOLECULAR MECHANICS CORRECTION TO CONH BONDS
MOLDAT   - PRINT DETAILS OF WORKING IN MOLDAT
MULLIK   - PRINT THE MULLIKEN POPULATION ANALYSIS
NLLSQ    - MINIMIZE GRADIENTS USING NLLSQ
NOINTER  - DO NOT PRINT INTERATOMIC DISTANCES
NOMM     - DO NOT USE MOLECULAR MECHANICS CORRECTION TO CONH BONDS
NOXYZ    - DO NOT PRINT CARTESIAN COORDINATES
OLDENS   - READ INITIAL DENSITY MATRIX OFF DISK
OPEN     - OPEN-SHELL RHF CALCULATION REQUESTED
PI       - RESOLVE DENSITY MATRIX INTO SIGMA AND PI BONDS
PL       - MONITOR CONVERGENCE OF DENSITY MATRIX IN ITER
PM3      - USE THE MNDO-PM3 HAMILTONIAN
POLAR    - CALCULATE FIRST, SECOND AND THIRD ORDER POLARIZABILITIES
POWSQ    - PRINT DETAILS OF WORKING IN POWSQ
PRECISE  - CRITERIA TO BE INCREASED BY 100 TIMES
PULAY    - USE PULAY'S CONVERGER TO OBTAIN A SCF
QUARTET  - RHF QUARTET STATE REQUIRED
QUINTET  - RHF QUINTET STATE REQUIRED
RESTART  - CALCULATION RESTARTED
ROOT=n   - ROOT n TO BE OPTIMIZED IN A C.I. CALCULATION
ROT=n    - THE SYMMETRY NUMBER OF THE SYSTEM IS n.
SADDLE   - OPTIMIZE TRANSITION STATE
SCFCRT=n - DEFAULT SCF CRITERION REPLACED BY THE VALUE SUPPLIED
SEARCH   - Use LINMIN instead of this keyword
SEXTET   - RHF SEXTET STATE REQUIRED
SHIFT=n  - A DAMPING FACTOR OF n DEFINED TO START SCF
SIGMA    - MINIMIZE GRADIENTS USING SIGMA
SINGLET  - RHF SINGLET STATE REQUIRED
SPIN     - PRINT FINAL UHF SPIN MATRIX
STEP1=n  - STEP SIZE n FOR FIRST COORDINATE IN GRID CALCULATION
STEP2=n  - STEP SIZE n FOR SECOND COORDINATE IN GRID CALCULATION
SYMMETRY - IMPOSE SYMMETRY CONDITIONS
T=n      - A TIME OF n SECONDS REQUESTED
THERMO   - PERFORM A THERMODYNAMICS CALCULATION
TIMES    - PRINT TIMES OF VARIOUS STAGES
T-PRIO   - TIME TAKES PRIORITY IN DRC
TRANS    - THE SYSTEM IS A TRANSITION STATE
           (USED IN THERMODYNAMICS CALCULATION)
TRIPLET  - TRIPLET STATE REQUIRED
UHF      - UNRESTRICTED HARTREE-FOCK CALCULATION
VECTORS  - PRINT FINAL EIGENVECTORS
VELOCITY - SUPPLY THE INITIAL VELOCITY VECTOR IN A DRC CALCULATION
X-PRIO   - GEOMETRY CHANGES TAKE PRIORITY IN DRC
XYZ      - DO ALL GEOMETRIC OPERATIONS IN CARTESIAN COORDINATES.
```

2.3   DEFINITIONS OF KEYWORDS

The definitions below are given with some technical expressions which are not further defined.   Interested users are referred to Appendix E of this manual to locate appropriate references which will provide further clarification.

There are three classes of keywords: (1) those which CONTROL substantial aspects of the calculation, i.e., those which affect the final heat of formation, (2) those which determine which OUTPUT will be calculated and printed, and (3) those which dictate the WORKING of the calculation, but which do not affect the heat of formation.   The assignment to one of these classes is designated by a (C), (O) or (W), respectively, following each keyword in the list below.

0SCF (O)

The data can be read in and output, but no actual calculation is performed when this keyword is used.   This is useful as a check on the input data to rule out errors introduced in transmission (usually a very last resort).

1ELECTRON (O)

The final one-electron matrix is printed out.   This matrix is composed of atomic orbitals; the array element between orbitals i and j on different atoms is given by

$$H(i,j) = 0.5 \times (beta(i) + beta(j)) \times overlap(i,j)$$

The matrix elements between orbitals i and j on the same atom are calculated from the electron-nuclear attraction energy, and also from the $U(i)$ value if i=j.

The one-electron matrix is unaffected by (a) the charge and (b) the electron density.   It is only a function of the geometry.   Abbreviation: 1ELEC.

1SCF (C)

When users want to examine the results of a single SCF calculation of a geometry, 1SCF should be used.   All the keywords relevant to output can be used.   If the gradients are to be calculated, then GRADIENTS should be specified as they are not calculated by default.

If the keyword RESTART is also present, then the geometric parameters which were being optimized will be used in the gradient calculation.

1SCF is helpful in a learning situation.   MOPAC normally performs many SCF calculations, and in order to minimize output when following the working of the SCF calculation, 1SCF is very useful.

ANALYT (W)

By default, finite difference derivatives of energy with respect to geometry are used. If ANALYT is specified, then analytical derivatives are used instead. Since the analytical derivatives are over Gaussian functions -- a STO-6G basis set is used -- the overlaps are also over Gaussian functions. This will result in a very small (less than 0.1 Kcal/mole) change in heat of formation. Use analytical derivatives (a) when the mantissa used is less than about 51-53 bits, or (b) when comparison with finite difference is desired. Finite difference derivatives are still used when non-variationally optimized wavefunctions are present.

AM1 (C)

The new AM1 method is to be used. By default MNDO is run.

BAR=n.nn (W)

In the SADDLE calculation the distance between the two geometries is steadily reduced until the transition state is located. Sometimes, however, the user may want to alter the maximum rate at which the distance between the two geometries reduces. BAR is a ratio, normally 0.15, or 15 percent. This represents a maximum rate of reduction of the bar of 15 percent per step. Alternative values that might be considered are BAR=0.05 or BAR=0.10, although other values may be used. See also SADDLE.

BIRADICAL (C)

NOTE: BIRADICAL is a redundant keyword, and represents a particular configuration interaction calculation. Experienced users of MECI (q.v.) can duplicate the effect of the keyword BIRADICAL by using the MECI keywords OPEN(2,2) and SINGLET.

For molecules which are believed to have biradicaloid character the option exists to optimize the lowest singlet energy state which results from the mixing of three states. These states are, in order, (1) the (micro)state arising from a one electron excitation from the HOMO to the LUMO, which is combined with the microstate resulting from the time-reversal operator acting on the parent microstate, the result being a full singlet state; (2) the state resulting from de-excitation from the formal LUMO to the HOMO; and (3) the state resulting from the single electron in the formal HOMO being excited into the LUMO.

	Microstate 1				Microstate 2		Microstate 3	
	Alpha	Beta	Alpha	Beta	Alpha	Beta	Alpha	Beta
LUMO	*		*				*	*
	---	---	---	---	---	---	---	---
			+					
HOMO		*	*		*	*		
	---	---	---	---	---	---	---	---

A configuration interaction calculation is involved here. A biradical calculation done without C.I. at the RHF level would be meaningless. Either rotational invariance would be lost, as in the D2d form of ethylene, or very artificial barriers to rotations would be found, such as in a methane molecule "orbiting" a D2d ethylene. In both cases the inclusion of limited configuration interaction corrects the error. BIRADICAL should not be used if either the HOMO or LUMO is degenerate; in this case, the full manifold of HOMO x LUMO should be included in the C.I., using MECI options. The user should be aware of this situation. When the biradical calculation is performed correctly, the result is normally a net stabilization. However, if the first singlet excited state is much higher in energy than the closed-shell ground state, BIRADICAL can lead to a destabilization. Abbreviation: BIRAD. See also MECI, C.I., OPEN, SINGLET.

BONDS (O)

The rotationally invariant bond order between all pairs of atoms is printed. In this context a bond is defined as the sum of the squares of the density matrix elements connecting any two atoms. For ethane, ethylene, and acetylene the carbon-carbon bond orders are roughly 1.00, 2.00, and 3.00 respectively. The diagonal terms are the valencies calculated from the atomic terms only and are defined as the sum of the bonds the atom makes with other atoms. In UHF and non-variationally optimized wavefunctions the calculated valency will be incorrect, the degree of error being proportional to the non-duodempotency of the density matrix. For an RHF wavefunction the square of the density matrix is equal to twice the density matrix.

The bonding contributions of all M.O.'s in the system are printed immediately before the bonds matrix. The idea of molecular orbital valency was developed by Gopinathan, Siddarth, and Ravimohan. Just as an atomic orbital has a 'valency', so has a molecular orbital. This leads to the following relations: The sum of the bonding contributions of all occupied M.O.'s is the same as the sum of all valencies which, in turn is equal to two times the sum of all bonds. The sum of the bonding contributions of all M.O.'s is zero.

C.I. (C)

Normally configuration interaction is invoked if any of the keywords which imply a C.I. calculation are used, such as BIRADICAL, TRIPLET or QUARTET. Note that ROOT= does not imply a C.I. calculation: ROOT= is only used when a C.I. calculation is done. However, as these implied C.I.'s involve the minimum number of configurations practical, the user may want to define a larger than minimum C.I., in which case the keyword C.I.=n can be used. When C.I.=n is specified, the n M.O.'s which "bracket" the occupied- virtual energy levels will be used. Thus, C.I.=2 will include both the HOMO and the LUMO, while C.I.=1 (implied for odd-electron systems) will only include the HOMO (This will do nothing for a closed-shell system, and leads to Dewar's half-electron correction for odd-electron systems). Users should be aware of the rapid increase in the size of the C.I. with increasing numbers of M.O.'s being used. Numbers of microstates implied by the use of the keyword C.I.=n on its own are as follows:

Keyword	Even-electron systems No. of electrons, configs			Odd-electron systems No. of electrons, configs		
	Alpha	Beta		Alpha	Beta	
C.I.=1	1	1	1	1	0	1
C.I.=2	1	1	4	1	0	2
C.I.=3	2	2	9	2	1	9
C.I.=4	2	2	36	2	1	24
C.I.=5	3	3	100	3	2	100
C.I.=6	3	3	400	3	2	300
C.I.=7	4	4	1225	4	3	1225
C.I.=8	(Do not use unless other keywords also used, see below)					

If a change of spin is defined, then larger numbers of M.O.'s can be used up to a maximum of 10. The C.I. matrix is of size 100 x 100. For calculations involving up to 100 configurations, the spin-states are exact eigenstates of the spin operators. For systems with more than 100 configurations, the 100 configurations of lowest energy are used. See also MICROS and the keywords defining spin-states.

Note that for any system, use of C.I.=5 or higher normally implies the diagonalization of a 100 by 100 matrix. As a geometry optimization using a C.I. requires the derivatives to be calculated using full SCF calculations, geometry optimization with large C.I.'s will require a considerable amount of time.

Associated keywords: MECI, ROOT=, SINGLET, DOUBLET, etc.

CHARGE=n (C)

When the system being studied is an ion, the charge, n, on the ion must be supplied by CHARGE=n. For cations n can be 1 or 2 or 3, etc, for anions -1 or -2 or -3, etc.

EXAMPLES

ION	KEYWORD	ION	KEYWORD
NH4(+)	CHARGE=1	CH3COO(-)	CHARGE=-1
C2H5(+)	CHARGE=1	(COO)(=)	CHARGE=-2
SO4(=)	CHARGE=-2	PO4(3-)	CHARGE=-3
HSO4(-)	CHARGE=-1	H2PO4(-)	CHARGE=-1

CYCLES=n (C)

In Bartel's method of gradient norm minimization, NLLSQ, the default number of cycles (100) is replaced by the number n specified by CYCLES=n.

DCART (O)

The cartesian derivatives which are calculated in DCART for variationally optimized systems are printed if the keyword DCART is present. The derivatives are in units of kcals/Angstrom, and the coordinates are displacements in x, y, and z.

DEBUG (O)

Certain keywords have specific output control meanings, such as FOCK, VECTORS and DENSITY. If they are used, only the final arrays of the relevant type are printed. If DEBUG is supplied, then all arrays are printed.   This is useful in debugging ITER. DEBUG can also increase the amount of output produced when certain output keywords are used, e.g. COMPFG.

DENOUT (O)

The density matrix at the end of the calculation is to be output in a form suitable for input in another job. If an automatic dump due to the time being exceeded occurs during the current run then DENOUT is invoked automatically. (see RESTART)

DENSITY (O)

At the end of a job, when the results are being printed, the density matrix is also printed. For RHF the normal density matrix is printed. For UHF the sum of the alpha and beta density matrices is printed.

If density is not requested, then the diagonal of the density matrix, i.e., the electron density on the atomic orbitals, will be printed.

DEP (O)

For use only with EXTERNAL=. When new parameters are published, they can be entered at run-time by using EXTERNAL=, but as this is somewhat clumsy, a permanent change can be made by use of DEP.

If DEP is invoked, a complete block of FORTRAN code will be generated, and this can be inserted directly into the BLOCK DATA file.

Note that this is designed only for use with MNDO or AM1 parameters. Only code for AM1 will be generated. To convert the FORTRAN code to define MNDO parameters, insert the letter M before every left parenthesis; thus, convert "(" to read "M(".

DEPVAR=n.nn (C)

In polymers the translation vector is frequently a multiple of some internal distance. For example, in polythene it is the C1-C3 distance. If a cluster unit cell of C6H12 is used, then symmetry can be used to tie together all the carbon atom coordinates and the translation vector distance. In this example DEPVAR=3.0 would be suitable.

DFP (W)

By default the Broyden-Fletcher-Goldfarb-Shanno method will be used to optimize geometries. The older Davidon-Fletcher-Powell method can be invoked by specifying DFP. This is intended to be used for comparison of the two methods.

DOUBLET (C)

When a configuration interaction calculation is done, all spin states are calculated simultaneously, either for component of spin = 0 or 1/2. When only doublet states are of interest, then DOUBLET can be specified, and all other spin states, while calculated, are ignored in the choice of root to be used.

Note that while almost every odd-electron system will have a doublet ground state, DOUBLET should still be specified if the desired state must be a doublet.

DOUBLET has no meaning in a UHF calculation.

DRC (C)

A Dynamic Reaction Coordinate calculation is to be run. By default, total energy is conserved, so that as the "reaction" proceeds in time, energy is transferred between kinetic and potential forms.

DRC=n.nnn (C)

In a DRC calculation, the "half-life" for loss of kinetic energy is defined as n.nnn x 10 femtoseconds. If n.nnn is set to zero, infinite damping simulating a very condensed phase is obtained.

This keyword cannot be written with spaces around the '=' sign.

DUMP (W)

Restart files are written automatically at one hour cpu time intervals to allow a long job to be restarted if the job is terminated catastrophically. To change the frequency of dump, set DUMP=nn to request a dump every nn seconds. Alternative form, DUMP=nnM for a dump every nn minutes. DUMP only works with geometry optimization, gradient minimization, and FORCE calculations. It does not (yet) work with a path or SADDLE calculation.

ECHO (O)

Data are echoed back if ECHO is specified. Only useful if data are suspected to be corrupt.

ENPART (O)

This is a very useful tool for analyzing the energy terms within a system. The total energy, in eV, obtained by the addition of the electronic and nuclear terms, is partitioned into mono- and bi-centric contributions, and these contributions in turn are divided into nuclear and one- and two-electron terms.

ESR (O)

The unpaired spin density arising from an odd-electron system can be calculated both RHF and UHF. In a UHF calculation the alpha and beta M.O.'s have different spatial forms, so unpaired spin density can naturally be present on in-plane hydrogen atoms such as in the phenoxy radical.

In the RHF formalism a MECI calculation is performed. If the keywords OPEN and C.I.= are both absent then only a single state is calculated. The unpaired spin density is then calculated from the state function. In order to have unpaired spin density on the hydrogens in, for example, the phenoxy radical, several states should be mixed.

EXCITED (C)

The state to be calculated is the first excited open-shell singlet state. If the ground state is a singlet, then the state calculated will be S(1); if the ground state is a triplet, then S(2). This state would normally be the state resulting from a one-electron excitation from the HOMO to the LUMO. Exceptions would be if the lowest singlet state were a biradical, in which case the EXCITED state could be a closed shell.

The EXCITED state will be calculated from a BIRADICAL calculation in which the second root of the C.I. matrix is selected. Note that the eigenvector of the C.I. matrix is not used in the current formalism. Abbreviation: EXCI.

NOTE: EXCITED is a redundant keyword, and represents a particular configuration interaction calculation. Experienced users of MECI can duplicate the effect of the keyword EXCITED by using the MECI keywords OPEN(2,2), SINGLET, and ROOT=2.

EXTERNAL=name (C)

Normally, AM1 and MNDO parameters are taken from the BLOCK DATA files within MOPAC. When the supplied parameters are not suitable, as in an element recently parameterized, and the parameters not yet installed in the user's copy of MOPAC, then the new parameters can be inserted at run time by use of EXTERNAL=<filename>, where <filename> is the name of the file which contains the new parameters.

<filename> consists of a series of parameter definitions in the format

<Parameter> <Element> <Value of parameter>

where the possible parameters are USS, UPP, UDD, ZS, ZP, ZD, BETAS, BETAP, BETAD, GSS, GSP, GPP, GP2, HSP, ALP, FNnm, n=1,2, or 3, and m=1 to 10, and the elements are defined by their chemical symbols, such as Si or SI.

When new parameters for elements are published, they can be typed in as shown. This file is ended by a blank line, the word END or nothing, i.e., no end-of-file delimiter. An example of a parameter data file would be:
Start of line| (Put at least 2 spaces before and after parameter name)

```
Line  1:     USS      Si      -34.08201495
Line  2:     UPP      Si      -28.03211675
Line  3:     BETAS    Si       -5.01104521
Line  4:     BETAP    Si       -2.23153969
Line  5:     ZS       Si        1.28184511
Line  6:     ZP       Si        1.84073175
Line  7:     ALP      Si        2.18688712
Line  8:     GSS      Si        9.82
Line  9:     GPP      Si        7.31
Line 10:     GSP      Si        8.36
Line 11:     GP2      Si        6.54
Line 12:     HSP      Si        1.32
```

Derived parameters do no need to be entered; they will be calculated from the optimized parameters. All "constants" such as the experimental heat of atomization are already inserted for all elements.

NOTE: EXTERNAL can only be used to input parameters for MNDO or AM1. It is unlikely, however, that any more MINDO/3 parameters will be published.

See also DEP to make a permanent change.

FILL=n (C)

The n'th M.O. in an RHF calculation is constrained to be filled. It has no effect on a UHF calculation. After the first iteration (NOTE: not after the first SCF calculation, but after the first iteration within the first SCF calculation) the n'th M.O. is stored, and, if occupied, no further action is taken at that time. If unoccupied, then the HOMO and the n'th M.O.'s are swapped around, so that the n'th M.O. is now filled. On all subsequent iterations the M.O. nearest in character to the stored M.O. is forced to be filled, and the stored M.O. replaced by that M.O. This is necessitated by the fact that in a reaction a particular M.O. may change its character very considerably. A useful procedure is to run 1SCF and DENOUT first, in order to identify the M.O.'s; the complete job is then run with OLDENS and FILL=nn, so that the eigenvectors at the first iteration are fully known. As FILL is known to give difficulty at times, consider also using C.I.=n and ROOT=m.

FLEPO (O)

The predicted and actual changes in the geometry, the derivatives, and search direction for each geometry optimization cycle are printed. This is useful if there is any question regarding the efficiency of the geometry optimizer.

FMAT

Details of the construction of the Hessian matrix for the force calculation are to be printed.

FORCE (C)

A force-calculation is to be run. The Hessian, that is the matrix (in millidynes per Angstrom) of second derivatives of the energy with respect to displacements of all pairs of atoms in x, y, and z, is calculated. On diagonalization this gives the force constants for the molecule. The force matrix, weighted for isotopic masses, is then used for calculating the vibrational frequencies. The system can be characterized as a ground state or a transition state by the presence of five (for a linear system) or six eigenvalues which are very small (less than about 30 reciprocal centimeters). A transition state is further characterized by one, and exactly one, negative force constant.

A FORCE calculation is a prerequisite for a THERMO calculation.

Before a FORCE calculation is started, a check is made to ensure that a stationary point is being used. This check involves calculating the gradient norm (GNORM) and if it is significant, the GNORM will be reduced using NLLSQ (Bartel's method). All internal coordinates are optimized, and any symmetry constraints are ignored at this point. An implication of this is that if the specification of the geometry relies on any angles being exactly 180 or zero degrees, the calculation may fail.

The geometric definition supplied to FORCE should not rely on angles or dihedrals assuming exact values. (The test of exact linearity is sufficiently slack that most molecules that are linear, such as acetylene and but-2-yne, should not be stopped.) See also THERMO, LET, TRANS, ISOTOPE.

FULSCF (W)

In line-searches the option exists to require all energy evaluations to be done using full SCF calculations. Normally full SCF calculations are not carried out during a line search as the density matrix is normally not changing very fast. The only important exception is in non-variationally optimized wavefunctions, such as occur in half-electron or C.I. calculations. Note: FULSCF will cause all derivatives to be calculated by explicit SCF calculations when non-variationally optimized wavefunctions are used. (This was in earlier copies of MOPAC but not documented.)

GEO-OK (W)

Normally the program will stop with a warning message if two atoms are within 0.8 Angstroms of each other, or, more rarely, the BFGS routine has difficulty optimizing the geometry. GEO-OK will over-ride the job termination sequence, and allow the calculation to proceed. In practice, most jobs that terminate due to these checks contain errors in data, so caution should be exercised if GEO-OK is used. An important exception to this warning is when the system contains, or may give rise to, a Hydrogen molecule. GEO-OK will override other geometric safety checks such as the unstable gradient in a geometry optimization preventing reliable optimization.

See also the message "GRADIENTS OF OLD GEOMETRY, GNORM= nn.nnnn"

GNORM=n.nn (W)

The geometry optimization termination criteria in both gradient minimization and energy minimization can be over-ridden by specifying a gradient norm requirement. For example, GNORM=20 would allow the geometry optimization to exit as soon as the gradient norm dropped below 20.0, the default being 1.0. A GNORM=0.01 could be used to refine a geometry beyond the normal limits. WARNING: If a very small value is chosen, the geometry optimization procedures may not terminate in a reasonable time. A reasonable lower bound for GNORM is about 0.1.

GRADIENTS (O)

In a 1SCF calculation gradients are not calculated by default: in non-variationally optimized systems this would take an excessive time. GRADIENTS allows the gradients to be calculated. All gradients are then calculated, whether marked for calculation or not, and printed. An exception is when the 1SCF was used in conjunction with the keyword RESTART, in which case only the coordinates being optimized would have their gradients printed. Abbreviation: GRAD.

GRAPH (O)

Information needed to generate electron density contour maps can be written to a file by calling GRAPH. GRAPH first calls MULLIK in order to generate the inverse-square-root of the overlap matrix, which is required for the re-normalization of the eigenvectors. All data essential for the graphics package DENSITY are then output.

H-PRIORITY (O)

In a DRC calculation, results will be printed whenever the calculated heat of formation changes by 0.1 Kcal/mole. Abbreviation: H-PRIO.

H-PRIORITY=n.nn (O)

In a DRC calculation, results will be printed whenever the calculated heat of formation changes by n.nn Kcal/mole.

IRC (C)

An Intrinsic Reaction Coordinate calculation is to be run. All kinetic energy is shed at every point in the calculation. See Background.

IRC=n (C)

An Intrinsic Reaction Coordinate calculation to be run; an initial perturbation in the direction of normal coordinate n to be applied. If n is negative, then perturbation is reversed, i.e., initial motion is in the opposite direction to the normal coordinate. This keyword cannot be written with spaces around the '=' sign.

ISOTOPE (O)

Generation of the FORCE matrix is very time-consuming, and in isotopic substitution studies several vibrational calculations may be needed. To allow the frequencies to be calculated from the (constant) force matrix, ISOTOPE is used. When a FORCE calculation is completed, ISOTOPE will cause the force matrix to be stored, regardless of whether or not any intervening restarts have been made. To re-calculate the frequencies, etc. starting at the end of the force matrix calculation, specify RESTART.

The two keywords RESTART and ISOTOPE can be used together. For example, if a normal FORCE calculation runs for a long time, the user may want to divide it up into stages and save the final force matrix. Once ISOTOPE has been used, it does not need to be used on subsequent RESTART runs.

ITRY=NN (W)

The default maximum number of SCF iterations is 200. When this limit presents difficulty, ITRY=nn can be used to re-define it. For example, if ITRY=400 is used, the maximum number of iterations will be set to 400. ITRY should normally not be changed until all other means of obtaining a SCF have been exhausted, e.g. PULAY CAMP-KING etc.

KINETIC=n.nnn (C)

In a DRC calculation n.nnn Kcals/mole of excess kinetic energy is added to the system as soon as the kinetic energy builds up to 0.2 Kcal/mole. The excess energy is added to the velocity vector, without change of direction.

LINMIN (O)

There are two line-minimization routines in MOPAC, an energy minimization and a gradient norm minimization. LINMIN will output details of the line minimization used in a given job.

LARGE (O)

Most of the time the output invoked by keywords is sufficient. LARGE will cause less-commonly wanted, but still useful, output to be printed. Currently LARGE only applies to the MECI.

LET (W)

As MOPAC evolves, the meaning of LET is changing.

Now LET means essentially "I know what I'm doing, override safety checks".

Currently, LET has the following meanings

1. In a FORCE calculation, it means that the supplied geometry is to be used, even if the gradients are large.

2. In a geometry optimization, the specified GNORM is to be used, even if it is less than 0.0001.

3. In a POLAR calculation, the molecule is to be orientated along its principal moments of inertia before the calculation starts. LET will prevent this step being done.

LOCALIZE (O)

The occupied eigenvectors are transformed into a localized set of M.O.'s by a series of 2 by 2 rotations which maximize $<psi**4>$. The value of $1/<psi**4>$ is a direct measure of the number of centers involved in the M.O.. Thus the value of $1/<psi**4>$ is 2.0 for H2, 3.0 for a three-center bond and 1.0 for a lone pair. Higher degeneracies than allowed by point group theory are readily obtained. For example, benzene would give rise to a 6-fold degenerate C-H bond, a 6-fold degenerate C-C sigma bond and a three-fold degenerate C-C pi bond. In principle, there is no single step method to unambiguously obtain the most localized set of M.O.'s in systems where several canonical structures are possible, just as no simple method exists for finding the most stable conformer of some large compound. However, the localized bonds generated will normally be quite acceptable for routine applications. Abbreviation: LOCAL.

MECI (O)

At the end of the calculation details of the Multi Electron Configuration Interaction calculation are printed if MECI is specified. The state vectors can be printed by specifying VECTORS. The MECI calculation is either invoked automatically, or explicitly invoked by the use of the C.I.=n keyword.

MICROS=n (C)

The microstates used by MECI are normally generated by use of a permutation operator. When individually defined microstates are desired, then MICROS=n can be used, where n defines the number of microstates to be read in.

Format for Microstates

After the geometry data plus any symmetry data are read in, data defining each microstate is read in, using format 20I1, one microstate per line. The microstate data is preceded by the word "MICROS" on a line by itself. There is at present no mechanism for using MICROS with a reaction path.

For a system with n M.O.'s in the C.I. (use OPEN=(nl,n) or C.I.=n to do this), the populations of the n alpha M.O.'s are defined, followed by the n beta M.O.'s. Allowed occupancies are zero and one. For n=6 the closed-shell ground state would be defined as 111000111000, meaning one electron in each of the first three alpha M.O.'s, and one electron in each of the first three beta M.O.'s.

Users are warned that they are responsible for completing any spin manifolds. Thus while the state 111100110000 is a triplet state with component of spin = 1, the state 111000110100, while having a component of spin = 0 is neither a singlet nor a triplet. In order to complete the spin manifold the microstate 110100111000 must also be included.

If a manifold of spin states is not complete, then the eigenstates of the spin operator will not be quantized. When and only when 100 or fewer microstates are supplied, loss of spin quantization occurs.

There are two other limitations on possible microstates. First, the number of electrons in every microstate should be the same. If they differ, a warning message will be printed, and the calculation continued (but the results will almost certainly be nonsense). Second, the component of spin for every microstate must be the same, except for teaching purposes. Two microstates of different components of spin will have a zero matrix element connecting them. No warning will be given as this is a reasonable operation in a teaching situation. For example, if all states arising from two electrons in two levels are to be calculated, say for teaching Russel-Saunders coupling, then the following microstates would be used:

Microstate	No. of alpha, beta electrons		Ms	State
1100	2	0	1	Triplet
1010	1	1	0	Singlet
1001	1	1	0	Mixed
0110	1	1	0	Mixed
0101	1	1	0	Singlet
0011	0	2	-1	Triplet

Constraints on the space manifold are just as rigorous, but much easier to satisfy. If the energy levels are degenerate, then all components of a manifold of degenerate M.O.'s should be either included or excluded. If only some, but not all, components are used, the required degeneracy of the states will be missing.

As an example, for the tetrahedral methane cation, if the user supplies the microstates corresponding to a component of spin = 3/2, neglecting Jahn-Teller distortion, the minimum number of states that can be supplied is 90 = (6!/(1!*5!))*(6!/(4!*2!)).

While the total number of electrons should be the same for all microstates, this number does not need to be the same as the number of electrons supplied to the C.I.; thus in the example above, a cationic state could be 110000111000.

The format is defined as 20I1 so that spaces can be used for empty M.O.'s.

MINDO/3 (C)

The default Hamiltonian within MOPAC is MNDO, with the alternatives of AM1 and MINDO/3. To use the MINDO/3 Hamiltonian the keyword MINDO/3 should be used. Acceptable alternatives to the keyword MINDO/3 are MINDO and MINDO3.

MMOK (C)

If the system contains a peptide linkage, then MMOK will allow a molecular mechanics correction to be applied so that the barrier to rotation is increased (to 14.00 Kcal/mole in N-methyl acetamide).

MULLIK (O)

A full Mulliken Population analysis is to be done on the final RHF wavefunction. This involves the following steps:
    (1) The eigenvector matrix is divided by the square root
        of the overlap matrix, S.
    (2) The Coulson-type density matrix, P, is formed.
    (3) The overlap population is formed from P(i,j)*S(i,j).
    (4) Half the off-diagonals are added onto the diagonals.

NLLSQ (C)

The gradient norm is to be minimized by Bartel's method. This is a Non-Linear Least Squares gradient minimization routine. Gradient minimization will locate one of three possible points:

(a) A minimum in the energy surface. The gradient norm will go to zero, and the lowest five or six eigenvalues resulting from a FORCE calculation will be approximately zero.

(b) A transition state. The gradient norm will vanish, as in (a), but in this case the system is characterized by one, and only one, negative force constant.

(c) A local minimum in the gradient norm space. In this (normally unwanted) case the gradient norm is minimized, but does not go to zero. A FORCE calculation will not give the five or six zero eigenvalues characteristic of a stationary point. While normally undesirable, this is sometimes the only way to obtain a geometry. For instance, if a system is formed which cannot be characterized as an intermediate, and at the same time is not a transition state, but nonetheless has some chemical significance, then that state can be refined using NLLSQ.

NOINTER (O)

The interatomic distances are printed by default. If you do not want them to be printed, specify NOINTER. For big jobs this reduces the output file considerably.

NOMM (C)

All four semi-empirical methods underestimate the barrier to rotation of a peptide bond. A Molecular Mechanics correction has been added which increases the barrier in N-methyl acetamide to 14 Kcal/mole. If you do not want this correction, specify NOMM (NO Molecular Mechanics).

NOXYZ (O)

The cartesian coordinates are printed by default. If you do not want them to be printed, specify NOXYZ. For big jobs this reduces the output file considerably.

OLDENS (W)

A density matrix produced by an earlier run of MOPAC is to be used to start the current calculation. This can be used in attempts to obtain an SCF when a previous calculation ended successfully but a subsequent run failed to go SCF.

OPEN(nl,n2) (C)

The M.O. occupancy during the SCF calculation can be defined in terms of doubly occupied, empty, and fractionally occupied M.O.'s. The fractionally occupied M.O.'s are defined by OPEN(nl,n2), where nl = number of electrons in the open-shell manifold, and n2 = number of open-shell M.O.'s; nl must be in the range 0 to 2. OPEN(1,1) will be assumed for odd-electron systems unless an OPEN keyword is used. Errors introduced by use of fractional occupancy are automatically corrected in a MECI calculation when OPEN(nl,n2) is used.

PI (O)

The normal density matrix is composed of atomic orbitals, that is s, px, py and pz. PI allows the user to see how each atom-atom interaction is split into sigma and pi bonds. The resulting "density matrix" is composed of the following basis-functions:- s-sigma, p-sigma, p-pi, d-sigma, d-pi, d-dell. The on-diagonal terms give the hybridization state, so that an sp2 hybridized system would be represented as s-sigma 1.0, p-sigma 2.0, p-pi 1.0

PM3

The PM3 method is to be used. PM3 is still very new, and users are cautioned that as yet we do not know how well it works for transition states or for vibrational frequencies.

POLAR

The polarizability and first and second hyperpolarizabilities are to be calculated. At present this calculation does not work for polymers, but should work for all other systems.

POWSQ (C)

Details of the working of POWSQ are printed out. This is only useful in debugging.

PRECISE (W)

The criteria for terminating all optimizations, electronic and geometric, are to be increased by a factor, normally, 100. This can be used where more precise results are wanted. If the results are going to be used in a FORCE calculation, where the geometry needs to be known quite precisely, then PRECISE is recommended; for small systems the extra cost in CPU time is minimal.

PULAY (W)

The default converger in the SCF calculation is to be replaced by Pulay's procedure as soon as the density matrix is sufficiently stable. A considerable improvement in speed can be achieved by the use of PULAY. If a large number of SCF calculations are envisaged, a sample calculation using 1SCF and PULAY should be compared with using 1SCF on its own, and if a saving in time results, then PULAY should be used in the full calculation. PULAY should be used with care in that its use will prevent the combined package of convergers (SHIFT, PULAY and the CAMP-KING convergers) from automatically being used in the event that the system fails to go SCF in (ITRY-10) iterations.

The combined set of convergers very seldom fails.

QUARTET (C)

The desired spin-state is a quartet, i.e., the state with component of spin = 1/2 and spin = 3/2. When a configuration interaction calculation is done, all spin states of spin equal to, or greater than 1/2 are calculated simultaneously, for component of spin = 1/2. From these states the quartet states are selected when QUARTET is specified, and all other spin states, while calculated, are ignored in the choice of root to be used. If QUARTET is used on its own, then a single state, corresponding to an alpha electron in each of three M.O.'s is calculated.

QUARTET has no meaning in a UHF calculation.

QUINTET (C)

The desired spin-state is a quintet, that is, the state with component of spin = 0 and spin = 2. When a configuration interaction calculation is done, all spin states of spin equal to, or greater than 0 are calculated simultaneously, for component of spin = 0. From these states the quintet states are selected when QUINTET is specified, and the septet states, while calculated, will be ignored in the choice of root to be used. If QUINTET is used on its own, then a single state, corresponding to an alpha electron in each of four M.O.'s is calculated.

QUINTET has no meaning in a UHF calculation.

RESTART (W)

When a job has been stopped, for whatever reason, and intermediate results have been stored, then the calculation can be restarted at the point where it stopped by specifying RESTART. The most common cause of a job stopping before completion is its exceeding the time allocated. A saddle-point calculation has no restart, but the output file contains information which can easily be used to start the calculation from a point near to where it stopped.

It is not necessary to change the geometric data to reflect the new geometry, as a result the geometry printed at the start of a restarted job will be that of the original data, not that of the restarted file.

A convenient way to monitor a long run is to specify 1SCF and RESTART; this will give a normal output file at very little cost. NOTE 1: In the FORCE calculation two restarts are possible. These are (a) a restart in FLEPO if the geometry was not optimized fully before FORCE was called, and (b) the normal restart in the construction of the force matrix. If the restart is in FLEPO within FORCE then the keyword FORCE should be deleted, and the keyword RESTART used on its own. Forgetting this point is a frequent cause of failed jobs. NOTE 2: Two restarts also exist in the IRC calculation. If an IRC calculation stops while in the FORCE calculation, then a normal restart can be done. If the job stops while doing the IRC calculation itself then the keyword IRC=n should be changed to IRC, or it can be omitted if DRC is also specified. The absence of the string "IRC=" is used to indicate that the FORCE calculation was completed before the restart files were written.

ROOT=n (C)

The n'th root of a C.I. calculation is to be used in the calculation. If a keyword specifying the spin-state is also present, e.g. SINGLET or TRIPLET, then the n'th root of that state will be selected. Thus ROOT=3 and SINGLET will select the third singlet root. If ROOT=3 is used on its own, then the third root will be used, which may be a triplet, the third singlet, or the second singlet (the second root might be a triplet). In normal use, this keyword would not be used. It is retained for educational and research purposes. Unusual care should be exercised when ROOT= is specified.

ROT=n (C)

In the calculation of the rotational contributions to the thermodynamic quantities the symmetry number of the molecule must be supplied. The symmetry number of a point group is the number of equivalent positions attainable by pure rotations. No reflections or improper rotations are allowed. This number cannot be assumed by default, and may be affected by subtle modifications to the molecule, such as isotopic substitution. A list of the most important symmetry numbers follows:

```
----    TABLE OF SYMMETRY NUMBERS     ----


C1 CI CS     1      D2 D2D D2H  4       C(INF)V   1
C2 C2V C2H   2      D3 D3D D3H  6       D(INF)H   2
C3 C3V C3H   3      D4 D4D D4H  8       T TD      12
C4 C4V C4H   4      D6 D6D D6H  12      OH        24
C6 C6V C6H   6      S6          3
```

SADDLE (C)

The transition state in a simple chemical reaction is to be optimized. Extra data are required. After the first geometry, specifying the reactants, and any symmetry functions have been defined, the second geometry, specifying the products, is defined, using the same format as that of the first geometry.

SADDLE often fails to work successfully. Frequently this is due to equivalent dihedral angles in the reactant and product differing by about 360 degrees rather than zero degrees. As the choice of dihedral can be difficult, users should consider running this calculation with the keyword XYZ. There is normally no ambiguity in the definition of cartesian coordinates. See also BAR=.

Many of the bugs in SADDLE have been removed in this version. Use of the XYZ option is strongly recommended.

SCFCRT=n.nn (W)

The default SCF criterion is to be replaced by that defined by SCFCRT=. The SCF criterion can be varied from about 0.001 to 1.D-25, although numbers in the range 0.0001 to 1.D-14 will suffice for most applications. To find a suitable value 1SCF and various values of SCFCRT=n.nnn should be used; a SCFCRT which allows evaluation of the heat of formation to an acceptable precision can thus be found rapidly. An overly tight criterion can lead to failure to achieve a SCF, and consequent failure of the run.

SEXTET (C)

The desired spin-state is a sextet: the state with component of spin = 1/2 and spin = 5/2.

The sextet states are the highest spin states normally calculable using MOPAC in its unmodified form. If SEXTET is used on its own, then a single state, corresponding to one alpha electron in each of five M.O.'s, is calculated. If several sextets are to be calculated, say the second or third, then OPEN(n1,n2) should be used.

SEXTET has no meaning in a UHF calculation.

SHIFT=n.nn (W)

In an attempt to obtain an SCF by damping oscillations which slow down the convergence or prevent an SCF being achieved, the virtual M.O. energy levels are shifted up or down in energy by a shift technique. The principle is that if the virtual M.O.'s are changed in energy relative to the occupied set, then the polarizability of the occupied M.O.'s will change pro rata. Normally, oscillations are due to autoregenerative charge fluctuations.

The SHIFT method has been re-written so that the value of SHIFT changes automatically to give a critically-damped system. This can result in a positive or negative shift of the virtual M.O. energy levels. If a non-zero SHIFT is specified, it will be used to start the SHIFT technique, rather than the default 15eV. If SHIFT=0 is specified, the SHIFT technique will not be used unless normal convergence techniques fail and the automatic "ALL CONVERGERS..." message is produced.

SIGMA (C)

The McIver-Komornicki gradient norm minimization routines, POWSQ and SEARCH are to be used. These are very rapid routines, but do not work for all species. If the gradient norm is low, i.e., less than about 5 units, then SIGMA will probably work; in most cases, NLLSQ is recommended. SIGMA first calculates a quite accurate Hessian matrix, a slow step, then works out the direction of fastest decent, and searches along that direction until the gradient norm is minimized. The Hessian is then partially updated in light of the new gradients, and a fresh search direction found. Clearly, if the Hessian changes markedly as a result of the line-search, the update done will be inaccurate, and the new search direction will be faulty.

SIGMA should be avoided if at all possible when non-variationally optimized calculations are being done.

If the Hessian is suspected to be corrupt within SIGMA it will be automatically recalculated. This frequently speeds up the rate at which the transition state is located. If you do not want the Hessian to be reinitialized -- it is costly in CPU time -- specify LET on the keyword line.

SINGLET (C)

When a configuration interaction calculation is done, all spin states are calculated simultaneously, either for component of spin = 0 or 1/2. When only singlet states are of interest, then SINGLET can be specified, and all other spin states, while calculated, are ignored in the choice of root to be used.

Note that while almost every even-electron system will have a singlet ground state, SINGLET should still be specified if the desired state must be a singlet.

SINGLET has no meaning in a UHF calculation, but see also TRIPLET.

SPIN (O)

The spin matrix, defined as the difference between the alpha and beta density matrices, is to be printed. If the system has a closed-shell ground state, e.g. methane run UHF, the spin matrix will be null.

If SPIN is not requested in a UHF calculation, then the diagonal of the spin matrix, that is the spin density on the atomic orbitals, will be printed.

STEP1=n.nnn (C)

In a grid calculation the step size in degrees or Angstroms for the first of the two parameters is given by n.nnn. 11 steps in each direction are calculated, giving a total of 121 steps. The origin is in the center at position (6,6).

STEP2=n.nnn (C)

In a grid calculation the step size in degrees or Angstroms for the second of the two parameters is given by n.nnn.

SYMMETRY (C)

Symmetry data defining related bond lengths, angles and dihedrals can be included by supplying additional data after the geometry has been entered. If there are any other data, such as values for the reaction coordinates, or a second geometry, as required by SADDLE, then it would follow the symmetry data. Symmetry data are terminated by one blank line. For non-variationally optimized systems symmetry constraints can save a lot of time because many derivatives do not need to be calculated. At the same time, there is a risk that the geometry may be wrongly specified, e.g. if methane radical cation is defined as being tetrahedral, no indication that this is faulty will be given until a FORCE calculation is run. (This system undergoes spontaneous Jahn-Teller distortion.)

Usually a lower heat of formation can be obtained when SYMMETRY is specified. To see why, consider the geometry of benzene. If no assumptions are made regarding the geometry, then all the C-C bond lengths will be very slightly different, and the angles will be almost, but not quite 120 degrees. Fixing all angles at 120 degrees, dihedrals at 180 or 0 degrees, and only optimizing one C-C and one C-H bond-length will result in a 2-D optimization, and exact D6h symmetry. Any deformation from this symmetry must involve error, so by imposing symmetry some error is removed.

The layout of the symmetry data is:

<defining atom> <symmetry relation> <defined atom> <defined atom>,... where the numerical code for <symmetry relation> is given in the table of SYMMETRY FUNCTIONS below.

For example, ethane, with three independent variables, can be defined as

```
SYMMETRY
ETHANE, D3D                                      NA NB NC

  C   0.000000 0     0.000000 0     0.000000 0    0  0  0
  C   1.528853 1     0.000000 0     0.000000 0    1  0  0
  H   1.105161 1   110.240079 1     0.000000 0    2  1  0
  H   1.105161 0   110.240079 0   120.000000 0    2  1  3
  H   1.105161 0   110.240079 0   240.000000 0    2  1  3
  H   1.105161 0   110.240079 0    60.000000 0    1  2  3
  H   1.105161 0   110.240079 0   180.000000 0    1  2  3
  H   1.105161 0   110.240079 0   300.000000 0    1  2  3
  0   0.000000 0     0.000000 0     0.000000 0    0  0  0
     3,    1,    4,    5,    6,    7,    8,
     3,    2,    4,    5,    6,    7,    8,
```

Here atom 3, a hydrogen, is used to define the bond lengths (symmetry relation 1) of atoms 4,5,6,7 and 8 with the atoms they are specified to bond with in the NA column of the data file; similarly, its angle (symmetry relation 2) is used to define the bond-angle of atoms 4,5,6,7 and 8 with the two atoms specified in the NA and NB columns of the data file. The other angles are point-group symmetry defined as a multiple of 60 degrees.

Spaces, tabs or commas can be used to separate data. Note that only three parameters are marked to be optimized. The symmetry data can be the last line of the data file unless more data follows, in which case a blank line must be inserted after the symmetry data.

The full list of available symmetry relations is as follows:

SYMMETRY FUNCTIONS

\<Symmetry
 relation\>

```
    1       BOND LENGTH    IS SET EQUAL TO THE REFERENCE BOND LENGTH
    2       BOND ANGLE     IS SET EQUAL TO THE REFERENCE BOND ANGLE
    3       DIHEDRAL ANGLE IS SET EQUAL TO THE REFERENCE DIHEDRAL ANGLE
    4       DIHEDRAL ANGLE VARIES AS  90 DEGREES - REFERENCE DIHEDRAL
    5       DIHEDRAL ANGLE VARIES AS  90 DEGREES + REFERENCE DIHEDRAL
    6       DIHEDRAL ANGLE VARIES AS 120 DEGREES - REFERENCE DIHEDRAL
    7       DIHEDRAL ANGLE VARIES AS 120 DEGREES + REFERENCE DIHEDRAL
    8       DIHEDRAL ANGLE VARIES AS 180 DEGREES - REFERENCE DIHEDRAL
    9       DIHEDRAL ANGLE VARIES AS 180 DEGREES + REFERENCE DIHEDRAL
   10       DIHEDRAL ANGLE VARIES AS 240 DEGREES - REFERENCE DIHEDRAL
   11       DIHEDRAL ANGLE VARIES AS 240 DEGREES + REFERENCE DIHEDRAL
   12       DIHEDRAL ANGLE VARIES AS 270 DEGREES - REFERENCE DIHEDRAL
   13       DIHEDRAL ANGLE VARIES AS 270 DEGREES + REFERENCE DIHEDRAL
   14       DIHEDRAL ANGLE VARIES AS THE NEGATIVE OF THE REFERENCE
            DIHEDRAL
   15       BOND LENGTH VARIES AS HALF THE REFERENCE BOND LENGTH
   16       BOND ANGLE VARIES AS HALF THE REFERENCE BOND ANGLE
   17       BOND ANGLE VARIES AS 180 DEGREES - REFERENCE BOND ANGLE
   18       BOND LENGTH IS A MULTIPLE OF REFERENCE BOND-LENGTH
```

Function 18 is intended for use in polymers, in which the translation vector may be a multiple of some bond-length. 1,2,3 and 14 are most commonly used. Abbreviation: SYM.

SYMMETRY is not available for use with cartesian coordinates.

T= (W)

This is a facility to allow the program to shut down in an orderly manner on computers with execution time C.P.U. limits.

The total C.P.U. time allowed for the current job is limited to nn.nn seconds; by default this is one hour, i.e., 3600 seconds. If the next cycle of the calculation cannot be completed without running a risk of exceeding the assigned time the calculation will write a restart file and then stop. The safety margin is 100 percent; that is, to do another cycle, enough time to do at least two full cycles must remain.

Alternative specifications of the time are T=nn.nnM, this defines the time in minutes, T=nn.nnH, in hours, and T=nn.nnD, in days, for very long jobs. This keyword cannot be written with spaces around the '=' sign.

THERMO (O)

The thermodynamic quantities, internal energy, heat capacity, partition function, and entropy can be calculated for translation, rotation and vibrational degrees of freedom for a single temperature, or a range of temperatures. Special situations such as linear systems and transition states are accommodated. The approximations used in the THERMO calculation are invalid below 100K, and checking of the lower bound of the temperature range is done to prevent temperatures of less than 100K being used.

Another limitation, for which no checking is done, is that there should be no internal rotations. If any exist, they will not be recognized as such, and the calculated quantities will be too low as a result.

In order to use THERMO the keyword FORCE must also be specified, as well as the value for the symmetry number; this is given by ROT=n.

If THERMO is specified on its own, then the default values of the temperature range are assumed. This starts at 200K and increases in steps of 10 degrees to 400K. Three options exist for overriding the default temperature range. These are:

THERMO(nnn) (O)

The thermodynamic quantities for a 200 degree range of temperatures, starting at nnnK and with an interval of 10 degrees are to be calculated.

THERMO(nnn,mmm) (O)

The thermodynamic quantities for the temperature range limited by a lower bound of nnn Kelvin and an upper bound of mmm Kelvin, the step size being calculated in order to give approximately 20 points, and a reasonable value for the step. The size of the step in Kelvin degrees will be 1, 2, or 5, or a power of 10 times these numbers.

THERMO(nnn,mmm,lll) (O)

Same as for THERMO(nnn,mmm), only now the user can explicitly define the step size. The step size cannot be less than 1K.

T-PRIORITY (O)

In a DRC calculation, results will be printed whenever the calculated time changes by 0.1 femtoseconds. Abbreviation, T-PRIO.

T-PRIORITY=n.nn (O)

In a DRC calculation, results will be printed whenever the calculated time changes by n.nn femtoseconds.

TRANS (C)

The imaginary frequency due to the reaction vector in a transition state calculation must not be included in the thermochemical calculation. The number of genuine vibrations considered can be:

3N-5 for a linear ground state system,

3N-6 for a non-linear ground state system, or

3N-6 for a linear transition-state complex,

3N-7 for a non-linear transition-state complex.

This keyword must be used in conjunction with THERMO if a transition state is being calculated.

TRANS=n (C)

The facility exists to allow the THERMO calculation to handle systems with internal rotations. TRANS=n will remove the n lowest vibrations. Note that TRANS=1 is equivalent to TRANS on its own. For xylene, for example, TRANS=2 would be suitable.

This keyword cannot be written with spaces around the '=' sign.

TRIPLET (C)

The triplet state is defined. If the system has an odd number of electrons, an error message will be printed.

UHF interpretation.

The number of alpha electrons exceeds that of the beta electrons by 2. If TRIPLET is not specified, then the numbers of alpha and beta electrons are set equal. This does not necessarily correspond to a singlet.

RHF interpretation.

An RHF MECI calculation is performed to calculate the triplet state. If no other C.I. keywords are used, then only one state is calculated by default. The occupancy of the M.O.'s in the SCF calculation is defined as (...2,1,1,0,..), that is, one electron is put in each of the two highest occupied M.O.'s.

See keywords C.I.=n and OPEN(n1,n2).

UHF (C)

The unrestricted Hartree-Fock Hamiltonian is to be used.

VECTORS (O)

The eigenvectors are to be printed. In UHF calculations both alpha
and beta eigenvectors are printed; in all cases the full set, occupied
and virtual, are output. The eigenvectors are normalized to unity, that
is the sum of the squares of the coefficients is exactly one. If DEBUG
is specified, then ALL eigenvectors on every iteration of every SCF
calculation will be printed. This is useful in a learning context, but
would normally be very undesirable.

VELOCITY (C)

The user can supply the initial velocity vector to start a DRC
calculation. Limitations have to be imposed on the geometry in order for
this keyword to work. These are (a) the input geometry must be in
cartesian coordinates, (b) the first three atoms must not be coaxial, (c)
triatomic systems are not allowed (See geometry specification - triatomic
systems are in internal coordinates, by definition.)

Put the velocity vector after the geometry as three data per line,
representing the x, y, and z components of velocity for each atom. The
units of velocity are centimeters per second.

The velocity vector will be rotated so as to suit the final
cartesian coordinate orientation of the molecule.

X-PRIORITY (O)

In a DRC calculation, results will be printed whenever   the
calculated geometry changes by 0.05 Angstroms. The geometry change is
defined as the linear sum of the translation vectors of motion for all
atoms in the system. Abbreviation, X-PRIO.

X-PRIORITY=n.nn (O)

In a DRC calculation, results will be printed whenever   the
calculated geometry changes by n.nn Angstroms.

XYZ (W)

The SADDLE calculation quite often fails due to faulty definition of the second geometry because the dihedrals give a lot of difficulty. To make this option easier to use, XYZ was developed. A calculation using XYZ runs entirely in cartesian coordinates, thus eliminating the problems associated with dihedrals. The connectivity of the two systems can be different, but the numbering must be the same. Dummy atoms can be used; these will be removed at the start of the run. A new numbering system will be generated by the program, when necessary.

XYZ is also useful for removing dummy atoms from an internal coordinate file; use XYZ and 1SCF.

If a large ring system is being optimized, sometimes the closure is difficult, in which case XYZ will normally work.

Except for SADDLE, do not use XYZ by default: use it only when something goes wrong!

CHAPTER 3

GEOMETRY SPECIFICATION

FORMAT: The geometry is read in using essentially "Free-Format" of
FORTRAN-77.   In fact, a character input is used in order to accommodate
the chemical symbols, but the numeric data can be regarded as
"free-format".   This means that integers and real numbers can be
interspersed, numbers can be separated by one or more spaces, a tab
and/or by one comma.   If a number is not specified, its value is set to
zero.

The geometry can be defined in terms of either internal or cartesian
coordinates.

INTERNAL COORDINATE DEFINITION

For any one atom (i) this consists of an interatomic distance in
Angstroms from an already-defined atom (j), an interatomic angle in
degrees between atoms i and j and an already defined k, (k and j must be
different atoms), and finally a torsional angle in degrees between atoms
i, j, k, and an already defined atom l (l cannot be the same as k or j).
See also dihedral angle coherency.

Exceptions:

1. Atom 1 has no coordinates at all: this is the origin.

2. Atom 2 must be connected to atom 1 by an interatomic distance
   only.

3. Atom 3 can be connected to atom 1 or 2, and must make an angle
   with atom 2 or 1 (thus - 3-2-1 or 3-1-2); no dihedral is
   possible for atom 3.   By default, atom 3 is connected to atom 2.

3.1 CONSTRAINTS

1. Interatomic distances must be greater than zero. Zero Angstroms
    is acceptable only if the parameter is symmetry-related to
    another atom, and is the dependent function.

2. Angles must be in the range 0.0 to 180.0, inclusive. This
    constraint is for the benefit of the user only; negative angles
    are the result of errors in the construction of the geometry,
    and angles greater than 180 degrees are fruitful sources of
    errors in the dihedrals.

3. Dihedrals angles must be definable. If atom i makes a dihedral
    with atoms j, k, and l, and the three atoms j, k, and l are in a
    straight line, then the dihedral has no definable angle. During
    the calculation this constraint is checked continuously, and if
    atoms j, k, and l lie within 0.02 Angstroms of a straight line,
    the calculation will output an error message and then stop. Two
    exceptions to this constraint are:

    (a) if the angle is zero or 180 degrees, in which case the
    dihedral is not used.

    (b) if atoms j, k, and l lie in an exactly straight line
    (usually the result of a symmetry constraint), as in acetylene,
    acetonitrile, but-2-yne, etc.

If the exceptions are used, care must be taken to ensure that the
program does not violate these constraints during any optimizations or
during any calculations of derivatives - see also FORCE.

Conversion to Cartesian Coordinates

By definition, atom 1 is at the origin of cartesian coordinate space
-- be careful, however, if atom 1 is a dummy atom. Atom 2 is defined as
lying on the positive X axis -- for atom 2, Y=0 and Z=0. Atom 3 is in
the X-Y plane unless the angle 3-2-1 is exactly 0 or 180 degrees. Atom
4, 5, 6, etc. can lie anywhere in 3-D space.

CARTESIAN COORDINATE DEFINITION

A definition of geometry in cartesian coordinates consists of the
chemical symbol or atomic number, then the cartesian coordinates and
optimization flags but no connectivity.

MOPAC uses the lack of connectivity to indicate that cartesian
coordinates are to be used. A unique case is the triatomics for which
only internal coordinates are allowed. This is to avoid conflict of
definitions: the user does not need to define the connectivity of atom
2, and can elect to use the default connectivity for atom 3.   As a
result, a triatomic may have no explicit connectivity defined, the user
thus taking advantage of the default connectivity.   Since internal
coordinates are more commonly used than cartesian, the above choice was

made.

If the keyword XYZ is absent every coordinate must be marked for optimization.   If any coordinates are not to be optimized, the keyword XYZ must be present. The coordinates of all atoms, including atoms 1, 2 and 3 can be optimized.   Dummy atoms should not be used, for obvious reasons.

3.2 DEFINITION OF ELEMENTS AND ISOTOPES

Elements are defined in terms of their atomic numbers or their chemical symbols. Acceptable symbols for MNDO are:

```
                       *                      *
   1   3   5   6   7   8   9  11  13  14  15  16  17  19  24  30  32  35
   H  Li   B   C   N   O   F  Na  Al  Si   P   S  Cl   K  Cr  Zn  Ge  Br
      LI                      NA  AL  SI          CL      CR  ZN  GE  BR
.eb

                   +   *   *   *   *   *   o
  50  53  80  92  99 102 103 104 105 106 107
  Sn   I  Hg  Pb  Xx  Cb  ++   +  --   -  Tv
  SN      HG  PB  XX  CB                  TV
```

* These symbols do not refer to elements which have been parameterized.
+ This is the dummy atom for assisting with geometry specification.
o This is the translation vector for use with polymers.

Old parameters for some elements are available. These are provided to allow compatibility with earlier copies of MOPAC. To use these older parameters, use a keyword composed of the chemical symbol followed by the year of publication of the parameters. Keywords currently available: Si1978 S1978.
For AM1, acceptable symbols are

```
.bb
 1 5 6 7 8 9 16 17 35 53 99 102 103 104 105 106 107
 H B C N O F S  Cl Br  I Xx  Cb  ++   +  --   -  Tv
.eb
              CL BR     XX  CB                  TV
```

If users need to use other elements such as Si or P, they can be specified.   In that case MNDO-type atoms will be used. As the behavior of such systems is not well investigated, users are cautioned to exercise unusual care with such systems. To alert users to this situation, the keyword PARASOK is defined.

For PM3, acceptable symbols are

```
 1 6 7 8 9 13 14 15 16 17 35 53 99 102 103 104 105 106 107
 H C N O F Al Si  P  S Cl Br  I Xx  Cb  ++   +  --   -  Tv
```

```
        AL SI        CL BR      XX  CB                 TV

        Diatomics Parameterized within the MINDO/3 Formalism
        H    B    C    N    O    F   Si   P    S   Cl   A star (*) indicates
     --------------------------------------------------   that the atom-pair is
   H    *    *    *    *    *    *    *    *    *    *    parameterized within
   B    *    *    *    *    *    *                        MINDO/3.
   C    *    *    *    *    *    *    *    *    *    *
   N    *    *    *    *    *    *              *    *
   O    *    *    *    *    *    *              *    *
   F    *    *    *    *    *    *              *
  Si    *         *                   *
   P    *         *                        *         *
   S    *         *    *    *    *              *    *
  Cl    *         *    *    *                   *    *
```

Extra entities available to MNDO, MINDO/3, AM1 and PM3

+ A 100% ionic alkali metal.

++ A 100% ionic alkaline earth metal.

- A 100% ionic halogen-like atom

-- A 100% ionic group VI-like atom.

Cb A special type of monovalent atom

Elements 103, 104, 105, and 106 are the sparkles; elements 11 and 19 are sparkles tailored to look like the alkaline metal ions; Tv is the translation vector for polymer calculations. See "Full description of sparkles".

Element 102, symbol Cb, is designed to satisfy valency requirements of atoms for which some bonds are not completed. Thus in "solid" diamond the usual way to complete the normal valency in a cluster model is to use hydrogen atoms. This approach has the defect that the electronegativity of hydrogen is different from that of carbon. The "Capped bond" atom, Cb, is designed to satisfy these valency requirements without acquiring a net charge.

Cb behaves like a monovalent atom, with the exception that it can alter its electronegativity to achieve an exactly zero charge in whatever environment it finds itself. It is thus all things to all atoms. On bonding to hydrogen it behaves similar to a hydrogen atom. On bonding to fluorine it behaves like a very electronegative atom. If several capped bond atoms are used, each will behave independently. Thus if the two hydrogen atoms in formic acid were replaced by Cb's then each Cb would independently become electroneutral.

Capped bonds should not be optimized. They are still very new and not enough is known yet. A fixed bond-length of 1.7 A is recommended, if two Cb are on one atom, a contained angle of 109.471221 degrees is suggested, and if three Cb are on one atom, a contained dihedral of -120

degrees (note sign) should be used.

Element 99, or XX is known as a dummy atom, and is used in the definition of the geometry; it is deleted automatically from any cartesian coordinate geometry files. *Dummy atoms are pure mathematic points, and are useful in defining geometries;* for example, in ammonia the definition of C3v symmetry is facilitated by using one dummy atom and symmetry relating the three hydrogens to it.

Output normally only gives chemical symbols.

Isotopes are used in conjunction with chemical symbols. If no isotope is specified, the average isotopic mass is used, thus chlorine is 35.453. This is different from all previous versions of MOPAC, in which the most abundant isotope was used by default. This change is justified by the removal of any ambiguity in the choice of isotope. Also, the experimental vibrational spectra involve a mixture of isotopes. If a user wishes to specify any specific isotope it should immediately follow the chemical symbol (no space), e.g., H2, H2.0140, Cl3, or Cl3.00335.

The sparkles ++, +, --, and - have no mass; if they are to be used in a force calculation, then appropriate masses should be used.

Each internal coordinate is followed by an integer, to indicate the action to be taken.

Integer	Action
1	Optimize the internal coordinate.
0	Do not optimize the internal coordinate.
-1	Reaction coordinate, or grid index.

Remarks:

Only one reaction coordinate is allowed, but this can be made more versatile by the use of SYMMETRY. If a reaction coordinate is used, the values of the reaction coordinate should follow immediately after the geometry and any symmetry data. No terminator is required, and free-format-type input is acceptable.

If two "reaction coordinates" are used, then MOPAC assumes that the two-dimensional space in the region of the supplied geometry is to be mapped. The two dimensions to be mapped are in the plane defined by the "-1" labels. Step sizes in the two directions must be supplied using STEP1 and STEP2 on the keyword line.

Using internal coordinates, the first atom has three unoptimizable coordinates, the second atom two, (the bond-length can be optimized) and the third atom has one unoptimizable coordinate. None of these six unoptimizable coordinates at the start of the geometry should be marked for optimization. If any are so marked, a warning is given, but the calculation will continue.

In cartesian coordinates all parameters can be optimized.

3.3 EXAMPLES OF COORDINATE DEFINITIONS.

Two examples will be given. The first is formic acid, HCOOH, and is presented in the normal style with internal coordinates. This is followed by formaldehyde, presented in such a manner as to demonstrate as many different features of the geometry definition as possible.

```
 MINDO/3
 Formic acid
 Example of normal geometry definition
   O                                             Atom 1 needs no coordinates.
   C    1.20 1                                   Atom 2 bonds to atom 1.
   O    1.32 1   116.8 1    0.0 0   2 1 0        Atom 3 bonds to atom 2 and
                                                 makes an angle with atom 1.
   H    0.98 1   123.9 1    0.0 0   3 2 1        Atom 4 has a dihedral of 0.0
                                                 with atoms 3, 2 and 1.
   H    1.11 1   127.3 1  180.0 0   2 1 3
   O    0.00 0     0.0 0    0.0 0   0 0 0
```

Atom 2, a carbon, is bonded to oxygen by a bond-length of 1.20 Angstroms, and to atom 3, an oxygen, by a bond-length of 1.32 Angstroms. The O-C-O angle is 116.8 degrees. The first hydrogen is bonded to the hydroxyl oxygen and the second hydrogen is bonded to the carbon atom. The H-C-O-O dihedral angle is 180 degrees.
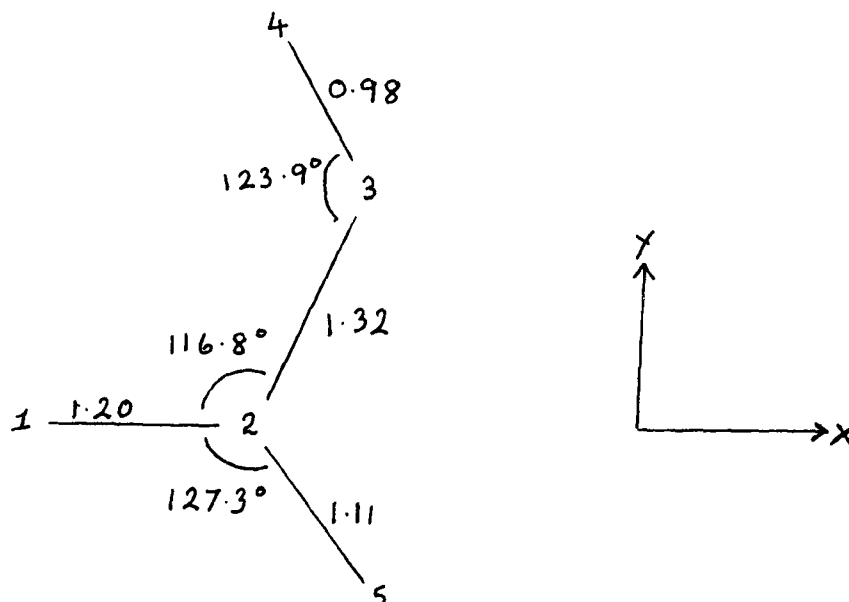
MOPAC can generate data-files, both in the Archive files, and at the end of the normal output file, when a job ends prematurely due to time restrictions. Here the coordinate definition for formic acid is shown.

Note that all coordinates are generated, as is the full connectivity. Also, the data are all neatly lined up. This is, of course, characteristic of machine-generated data, but is useful when checking for errors.

Format of internal coordinates in ARCHIVE file

O	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.209615	1	0.000000	0	0.000000	0	1	0	0
O	1.313679	1	116.886168	1	0.000000	0	2	1	0
H	0.964468	1	115.553316	1	0.000000	0	3	2	1
H	1.108040	1	128.726078	1	180.000000	0	2	1	3
O	0.000000	0	0.000000	0	0.000000	0	0	0	0

Polymers are defined by the presence of a translation vector. In the following example, polyethylene, the translation vector spans three monomeric units, and is 7.7 Angstroms long. Note in this example the presence of two dummy atoms. These not only make the geometry definition easier but also allow the translation vector to be specified in terms of distance only, rather than both distance and angles.

Example of polymer coordinates from ARCHIVE file

T=20000
  POLYETHYLENE, CLUSTER UNIT : C6H12

C	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.540714	1	0.000000	0	0.000000	0	1	0	0
C	1.542585	1	113.532306	1	0.000000	0	2	1	0
C	1.542988	1	113.373490	1	179.823613	1	3	2	1
C	1.545151	1	113.447508	1	179.811764	1	4	3	2
C	1.541777	1	113.859804	1	-179.862648	1	5	4	3
XX	1.542344	1	108.897076	1	-179.732346	1	6	5	4
XX	1.540749	1	108.360151	1	-178.950271	1	7	6	5
H	1.114786	1	90.070026	1	126.747447	1	1	3	2
H	1.114512	1	90.053136	1	-127.134856	1	1	3	2
H	1.114687	1	90.032722	1	126.717889	1	2	4	3
H	1.114748	1	89.975504	1	-127.034513	1	2	4	3
H	1.114474	1	90.063308	1	126.681098	1	3	5	4
H	1.114433	1	89.915262	1	-126.931090	1	3	5	4
H	1.114308	1	90.028131	1	127.007845	1	4	6	5
H	1.114434	1	90.189506	1	-126.759550	1	4	6	5
H	1.114534	1	88.522263	1	127.041363	1	5	7	6
H	1.114557	1	88.707407	1	-126.716355	1	5	7	6
H	1.114734	1	90.638631	1	127.793055	1	6	8	7
H	1.115150	1	91.747016	1	-126.187496	1	6	8	7
Tv	7.746928	1	0.000000	0	0.000000	0	1	7	8
O	0.000000	0	0.000000	0	0.000000	0	0	0	0

CHAPTER 4

EXAMPLES

In this chapter various examples of data-files are described. With
MOPAC comes two sets of data for running calculations. One of these is
called MNRSD1.DAT, and this will now be described.

4.1 MNRSD1 TEST DATA FILE FOR FORMALDEHYDE

The following file is suitable for generating the results described
in the next section, and would be suitable for debugging data.

```
Line  1:           SYMMETRY
Line  2:  Formaldehyde, for Demonstration Purposes
Line  3:
Line  4:  O
Line  5:  C 1.2 1
Line  6:  H 1.1 1 120 1
Line  7:  H 1.1 0 120 0 180 0 2 1 3
Line  8:
Line  9:  3 1 4
Line 10:  3 2 4
Line 11:
```

This data could be more neatly written as

```
Line  1:           SYMMETRY
Line  2:  Formaldehyde, for Demonstration Purposes
Line  3:
Line  4:  O   0.000000  0    0.000000  0    0.000000  0   0  0  0
Line  5:  C   1.200000  1    0.000000  0    0.000000  0   1  0  0
Line  6:  H   1.100000  1  120.000000  1    0.000000  0   2  1  0
Line  7:  H   1.100000  0  120.000000  0  180.000000  0   2  1  3
Line  8:  0   0.000000  0    0.000000  0    0.000000  0   0  0  0
Line  9:  3,   1,    4,
Line 10:  3,   2,    4,
Line 11:
```

These two data-files will produce identical results files.

In all geometric specifications care must be taken in defining the internal coordinates to ensure that no three atoms being used to define a fourth atom's dihedral angle ever fall into a straight line. This can happen in the course of a geometry optimization, in a SADDLE calculation or in following a reaction coordinate. If such a condition should develop, then the position of the dependent atom would become ill-defined.

4.2 MOPAC OUTPUT FOR TEST-DATA FILE MNRSD1

```
**************************************************************************
** FRANK J. SEILER RES. LAB., U.S. AIR FORCE ACADEMY, COLO. SPGS., CO. 80840 **
**************************************************************************
```

                    MNDO CALCULATION RESULTS   Note 1

```
**************************************************************************
*          MOPAC:  VERSION  5.00              CALC'D.  8-NOV-88    Note 2
*  SYMMETRY - SYMMETRY CONDITIONS TO BE IMPOSED
*   T=        - A TIME OF  3600.0 SECONDS REQUESTED
*  DUMP=N   - RESTART FILE WRITTEN EVERY  3600.0 SECONDS
****************************************************************043BY043
```

PARAMETER DEPENDENCE DATA

REFERENCE ATOM	FUNCTION NO.	DEPENDENT ATOM(S)
3	1	4
3	2	4

DESCRIPTIONS OF THE FUNCTIONS USED

| 1 | BOND LENGTH | IS SET EQUAL TO THE REFERENCE BOND LENGTH |
| 2 | BOND ANGLE | IS SET EQUAL TO THE REFERENCE BOND ANGLE |

SYMMETRY   Note 3

Formaldehyde, for Demonstration Purposes

ATOM NUMBER (I)	CHEMICAL SYMBOL	BOND LENGTH (ANGSTROMS) NA:I	BOND ANGLE (DEGREES) NB:NA:I	TWIST ANGLE (DEGREES) NC:NB:NA:I	NA	NB	NC
1	O					Note 4	
2	C	1.20000 *			1		
3	H	1.10000 *	120.00000 *		2	1	
4	H	1.10000	120.00000	180.00000	2	1	3

CARTESIAN COORDINATES

NO.	ATOM	X	Y	Z	
1	O	0.0000	0.0000	0.0000	
2	C	1.2000	0.0000	0.0000	Note 5
3	H	1.7500	0.9526	0.0000	
4	H	1.7500	-0.9526	0.0000	

H: (MNDO): M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4899, (1977)
C: (MNDO): M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4899, (1977)
O: (MNDO): M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4899, (1977)

RHF CALCULATION, NO. OF DOUBLY OCCUPIED LEVELS = 6

INTERATOMIC DISTANCES

0

		O 1	C 2	H 3	H 4	
		--------	--------	--------	--------	
O	1	0.000000				
C	2	1.200000	0.000000			
H	3	1.992486	1.100000	0.000000		Note 5
H	4	1.992486	1.100000	1.905256	0.000000	

CYCLE: 1 TIME: 1.96 TIME LEFT: 3596.3 GRAD.:   8.248 HEAT:-32.86456
CYCLE: 2 TIME: 0.93 TIME LEFT: 3595.3 GRAD.:   2.205 HEAT:-32.88052
HEAT OF FORMATION TEST SATISFIED   Note 7
PETERS TEST SATISFIED   Note 8

     SYMMETRY   Note 9
 Formaldehyde, for Demonstration Purposes   Note 10

     PETERS TEST WAS SATISFIED IN BFGS   OPTIMIZATION   Note 11
     SCF FIELD WAS ACHIEVED   Note 12

                         MNDO   CALCULATION   Note 13
                                                 VERSION 5.00
                                                 8-NOV-88

     FINAL HEAT OF FORMATION =   -32.88189 KCAL   Note 14

     TOTAL ENERGY   =   -478.11917 EV
     ELECTRONIC ENERGY   =   -870.73880 EV
     CORE-CORE REPULSION   =   392.61963 EV

     IONIZATION POTENTIAL =   11.04146
     NO. OF FILLED LEVELS =   6
     MOLECULAR WEIGHT   =   30.026

SCF CALCULATIONS =   15
COMPUTATION TIME =   6.280 SECONDS   Note 15

ATOM NUMBER (I)	CHEMICAL SYMBOL	BOND LENGTH (ANGSTROMS) NA:I	BOND ANGLE (DEGREES) NB:NA:I	TWIST ANGLE (DEGREES) NC:NB:NA:I	NA	NB	NC
1	O						
2	C	1.21646 *			1	Note 16	
3	H	1.10603 *	123.51031 *		2	1	
4	H	1.10603	123.51031	180.00000	2	1	3

INTERATOMIC DISTANCES

0

		O 1	C 2	H 3	H 4
O	1	0.000000			
C	2	1.216457	0.000000		
H	3	2.046624	1.106031	0.000000	
H	4	2.046624	1.106031	1.844387	0.000000

EIGENVALUES

-42.99217 -25.11966 -16.95495 -16.29956 -14.17982 -11.04146   0.85991   3.67635
   3.84965   7.12622   Note 17

NET ATOMIC CHARGES AND DIPOLE CONTRIBUTIONS

ATOM NO.	TYPE	CHARGE	ATOM ELECTRON DENSITY	
1	O	-0.2902	6.2902	
2	C	0.2921	3.7079	Note 18
3	H	-0.0010	1.0010	
4	H	-0.0010	1.0010	

DIPOLE	X	Y	Z	TOTAL	
POINT-CHG.	1.690	0.000	0.000	1.690	
HYBRID	0.475	0.000	0.000	0.475	Note 19
SUM	2.165	0.000	0.000	2.165	

CARTESIAN COORDINATES

NO.	ATOM	X	Y	Z
1	O	0.0000	0.0000	0.0000
2	C	1.2165	0.0000	0.0000
3	H	1.8271	0.9222	0.0000
4	H	1.8271	-0.9222	0.0000

ATOMIC ORBITAL ELECTRON POPULATIONS

| 1.88260 | 1.21603 | 1.89108 | 1.30048 | 1.25527 | 0.86214 | 0.89092 | 0.69952 |
| 1.00098 | 1.00098 | | | | | | Note 20 |

TOTAL CPU TIME: 6.48 SECONDS

== MOPAC DONE ==

NOTES ON RESULTS FILE

NOTE 1: The banner indicates whether the calculation uses a MNDO, MINDO/3, AM1 or PM3 Hamiltonian; here, the default MNDO Hamiltonian is used.

NOTE 2: The Version number is a constant for any release of MOPAC, and refers to the program, not to the Hamiltonians used. The version number should be cited in any correspondence regarding MOPAC.   Users' own in-house modified versions of MOPAC will have a final digit different from zero, e.g. 5.01.

All the keywords used, along with a brief explanation, should be printed at this time.   If a keyword is not printed, it has not been recognized by the program. Keywords can be in upper or lower case letters, or any mixture. The cryptic message at the right end of the lower line of asterisks indicates the number of heavy and light atoms this version of MOPAC is configured for.

NOTE 3: Symmetry information is output to allow the user to verify that the   requested symmetry functions have in fact been recognized and used.

NOTE 4: The data for this example used a mixture of atomic numbers and chemical symbols, but the internal coordinate output is consistently in chemical symbols.

The atoms in the system are, in order:

Atom 1, an oxygen atom; this is defined as being at the origin.

Atom 2, the carbon atom. Defined as being 1.2 Angstroms from the oxygen atom, it is located in the +x direction. This distance is marked for optimization.

Atom 3, a hydrogen atom.   It is defined as being 1.1 Angstroms from the carbon atom, and making an angle of 120 degrees with the oxygen atom. The asterisks indicate that the bond length and angle are both to be optimized.

Atom 4, a hydrogen atom. The bond length supplied has been overwritten with the symmetry-defined C-H bond length. Atom 4 is defined as being 1.1 Angstroms from atom 2, making a bond-angle

of 120 degrees with atom 1, and a dihedral angle of 180 degrees with atom 3.

None of the coordinates of atom 4 are marked for optimization. The bond-length and angle are symmetry-defined by atom 3, and the dihedral is group-theory symmetry-defined as being 180 degrees. (The molecule is flat.)

NOTE 5: The cartesian coordinates are calculated as follows:

Stage 1: The coordinate of the first atom is defined as being at the origin of cartesian space, while the coordinate of the second atom is defined as being displaced by its defined bond length along the positive x-axis. The coordinate of the third atom is defined as being displaced by its bond length in the x-y plane, from either atom 1 or 2 as defined in the data, or from atom 2 if no numbering is given. The angle it makes with atoms 1 and 2 is that given by its bond angle.

The dihedral, which first appears in the fourth atom, is defined according to the I.U.P.A.C. convention. NOTE: This is different from previous versions of MNDO and MINDO/3, where the dihedral had the opposite chirality to that defined by the I.U.P.A.C. convention.

Stage 2: Any dummy atoms are removed. As this particular system contains no dummy atoms, nothing is done.

NOTE 6: The interatomic distances are output for the user's advice, and a simple check made to insure that the smallest interatomic distance is greater than 0.8 Angstroms.

NOTE 7: The geometry is optimized in a series of cycles, each cycle consisting of a line search and calculation of the gradients. The time given is the C.P.U. time for the cycle; time left is the total time requested (here 100 seconds) less the C.P.U. time since the start of the calculation (which is earlier than the start of the first cycle!). These times can vary slightly from cycle to cycle due to different options being used, for example whether or not two or more SCF calculations need to be done to ensure that the heat of formation is lowered. The gradient is the scalar length in kcal/mole/Angstrom of the gradient vector.

NOTE 8: At the end of the BFGS geometry optimization a message is given which indicates how the optimization ended. All "normal" termination messages contain the word "satisfied"; other terminations may give acceptable results, but more care should be taken, particularly regarding the gradient vector.

NOTE 9 and 10: The keywords used, titles and comments are reproduced here to remind the user of the name of the calculation.

NOTE 11 and 12: Two messages are given here. The first is a reminder of how the geometry was obtained, whether from the Broyden-Fletcher-Goldfarb-Shanno, Bartel's or the McIver-Komornicki

methods.   For any further results to be printed the second message must
be as shown; when no SCF is obtained no results will be printed.

NOTE 13:   Again, the results are headed with either MNDO or MINDO/3
banners,   and   the version number.   The date has been moved to below the
version number for convenience.

NOTE 14:   The total energy of the system is   the   addition   of   the
electronic   and nuclear terms.   The heat of formation is relative to the
elements in their standard state.   The I.P.   is   the   negative   of   the
energy   level   of   the   highest   occupied, or highest partially occupied
molecular orbital (in accordance with Koopmans' theorem).

NOTE 15:   Advice on time required for   the   calculation.   This   is
obviously useful in estimating the times required for other systems.

NOTE 16:   The fully optimized geometry   is   printed   here.   If   a
parameter   is not marked for optimization, it will not be changed unless
it is a symmetry-related parameter.

NOTE 17:   The   roots   are   the   eigenvalues   or   energy   levels   in
electron   volts of the molecular orbitals.   There are six filled levels,
therefore   the   HOMO   has   an   energy   of   -11.041eV;   analysis   of   the
corresponding   eigenvector   (not   given   here)   shows   that it is mainly
lone-pair on oxygen.   The eigenvectors form an orthonormal set.

NOTE 18:   The charge on an atom is the sum   of   the   positive   core
charge; for hydrogen, carbon, and oxygen these numbers are 1.0, 4.0, and
6.0, respectively, and the negative of the number of valence   electrons,
or   atom   electron   density on the atom, here 1.0010, 3.7079, and 6.2902
respectively.

NOTE 19:   The   dipole   is   the   scalar   of   the   dipole   vector   in
cartesian   coordinates.   The   components of the vector coefficients are
the point-charge dipole and the hybridization dipole.   In   formaldehyde
there is no z-dipole since the molecule is flat.

NOTE 20:   MNDO AM1, PM3, and MINDO/3 all use   the   Coulson   density
matrix.   Only   the   diagonal   elements   of the matrix, representing the
valence orbital   electron   populations,   will   be   printed,   unless   the
keyword DENSITY is specified.

Two extra lines are added as a result of user requests:

(1) The total CPU time   for   the   job,   excluding   loading   of   the
executable, is printed.

(2) In order to know that MOPAC has ended,   the   message   ==   MOPAC
DONE == is printed.

This example is taken from the first data-file in TESTDATA.DAT, and illustrates the working of a FORCE calculation.

5.1 DATA FILE FOR A FORCE CALCULATION

```
Line 1 : ROT=2 THERMO(298,298,) PRECISE FORCE ISOTOPE SYMMETRY
Line 2 :    DEMONSTRATION OF MOPAC - FORCE AND THERMODYNAMICS CALCULATION
Line 3 :    FORMALDEHYDE, MNDO ENERGY = -32.8819
Line 4a:    O    0.000000  0    0.000000  0    0.000000  0   0  0  0
Line 4b:    C    1.216487  1    0.000000  0    0.000000  0   1  0  0
Line 4c:    H    1.106109  1  123.513310  1    0.000000  0   2  1  0
Line 4d:    H    1.106109  1  123.513310  1  180.000000  1   2  1  3
Line 4e:    0    0.000000  0    0.000000  0    0.000000  0   0  0  0
Line 5a:    3,   1,   4,
Line 5b:    3,   2,   4,
```

5.2 RESULTS FILE FOR THE FORCE CALCULATION

```
****************************************************************************
** FRANK J. SEILER RES. LAB., U.S. AIR FORCE ACADEMY, COLO. SPGS., CO. 80840 **
****************************************************************************

                      MNDO CALCULATION RESULTS


****************************************************************************
*           MOPAC:  VERSION  5.00              CALC'D.  8-NOV-88
*   SYMMETRY - SYMMETRY CONDITIONS TO BE IMPOSED
*   T=       - A TIME OF  3600.0 SECONDS REQUESTED
*   DUMP=N   - RESTART FILE WRITTEN EVERY  3600.0 SECONDS
*   FORCE    - FORCE CALCULATION SPECIFIED                      Note 1
*   PRECISE  - CRITERIA TO BE INCREASED BY 100 TIMES
*   ISOTOPE  - FORCE MATRIX WRITTEN TO DISK (CHAN. 9 )
```

* THERMO   - THERMODYNAMIC QUANTITIES TO BE CALCULATED
* ROT   - SYMMETRY NUMBER OF 2 SPECIFIED
**043BY043

PARAMETER DEPENDENCE DATA

REFERENCE ATOM	FUNCTION NO.	DEPENDENT ATOM(S)
3	1	4
3	2	4

DESCRIPTIONS OF THE FUNCTIONS USED

1   BOND LENGTH   IS SET EQUAL TO THE REFERENCE BOND LENGTH
2   BOND ANGLE   IS SET EQUAL TO THE REFERENCE BOND ANGLE
ROT=2 THERMO(298,298,) PRECISE FORCE ISOTOPE SYMMETRY
DEMONSTRATION OF MOPAC - FORCE AND THERMODYNAMICS CALCULATION
FORMALDEHYDE, MNDO ENERGY = -32.8819

ATOM NUMBER (I)	CHEMICAL SYMBOL	BOND LENGTH (ANGSTROMS) NA:I	BOND ANGLE (DEGREES) NB:NA:I	TWIST ANGLE (DEGREES) NC:NB:NA:I	NA	NB	NC
1	O						
2	C	1.21649 *			1		
3	H	1.10611 *	123.51331 *		2	1	
4	H	1.10611 *	123.51331 *	180.00000 *	2	1	3

CARTESIAN COORDINATES

NO.	ATOM	X	Y	Z
1	O	0.0000	0.0000	0.0000
2	C	1.2165	0.0000	0.0000
3	H	1.8272	0.9222	0.0000
4	H	1.8272	-0.9222	0.0000

H: (MNDO): M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4899, (1977)
C: (MNDO): M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4899, (1977)
O: (MNDO): M.J.S. DEWAR, W. THIEL, J. AM. CHEM. SOC., 99, 4899, (1977)

RHF CALCULATION, NO. OF DOUBLY OCCUPIED LEVELS = 6

INTERATOMIC DISTANCES

0

		O 1	C 2	H 3	H 4
O	1	0.000000			
C	2	1.216487	0.000000		
H	3	2.046748	1.106109	0.000000	
H	4	2.046748	1.106109	1.844454	0.000000

HEAT OF FORMATION = -32.881900 KCALS/MOLE

INTERNAL COORDINATE DERIVATIVES

ATOM	AT. NO.	BOND	ANGLE	DIHEDRAL
1	O			
2	C	0.000682		
3	H	0.000223	-0.000140	
4	H	0.000000	0.000000	0.000000

Note 2

GRADIENT NORM =   0.00073

TIME FOR SCF CALCULATION =   1.29

TIME FOR DERIVATIVES   =   0.49   Note 3

SYMMETRY WAS SPECIFIED, BUT CANNOT BE USED HERE

MOLECULAR WEIGHT =   30.03

PRINCIPAL MOMENTS OF INERTIA IN CM(-1)

A =   9.832732   B =   1.261998   C =   1.118449

PRINCIPAL MOMENTS OF INERTIA IN UNITS OF 10**(-40)*GRAM-CM**2

A =   2.846883   B =   22.181200   C =   25.028083

ORIENTATION OF MOLECULE IN FORCE CALCULATION

NO.	ATOM	X	Y	Z
1	8	-0.6093	0.0000	0.0000
2	6	0.6072	0.0000	0.0000
3	1	1.2179	0.9222	0.0000
4	1	1.2179	-0.9222	0.0000

FIRST DERIVATIVES WILL BE USED IN THE CALCULATION OF SECOND DERIVATIVES

DEFAULT TIME OF 3600.00 SECONDS ALLOCATED FOR THIS STEP

ESTIMATED TIME TO COMPLETE CALCULATION =   85.44 SECONDS
STEP: 1 TIME =   3.78 SECS, INTEGRAL =   3.78 TIME LEFT: 3594.44

- 56 -

```
STEP:   2 TIME =      3.97 SECS, INTEGRAL =      7.75 TIME LEFT:   3590.47
STEP:   3 TIME =      4.12 SECS, INTEGRAL =     11.87 TIME LEFT:   3586.35
STEP:   4 TIME =      3.81 SECS, INTEGRAL =     15.68 TIME LEFT:   3582.54
STEP:   5 TIME =      3.95 SECS, INTEGRAL =     19.63 TIME LEFT:   3578.59
STEP:   6 TIME =      3.98 SECS, INTEGRAL =     23.61 TIME LEFT:   3574.61
STEP:   7 TIME =      3.74 SECS, INTEGRAL =     27.35 TIME LEFT:   3570.87
STEP:   8 TIME =      3.68 SECS, INTEGRAL =     31.03 TIME LEFT:   3567.19
STEP:   9 TIME =      5.17 SECS, INTEGRAL =     36.20 TIME LEFT:   3562.02
STEP:  10 TIME =      3.96 SECS, INTEGRAL =     40.16 TIME LEFT:   3558.06
STEP:  11 TIME =      3.70 SECS, INTEGRAL =     43.86 TIME LEFT:   3554.36
STEP:  12 TIME =      5.08 SECS, INTEGRAL =     48.94 TIME LEFT:   3549.28
```

FORCE MATRIX IN MILLIDYNES/ANGSTROM

0

		O 1	C 2	H 3	H 4
O	1	9.557484			
C	2	8.682897	11.426636		
H	3	0.598872	2.553314	3.034863	
H	4	0.598872	2.553316	0.304427	3.034864

HEAT OF FORMATION = -32.881900 KCALS/MOLE

ZERO POINT ENERGY   18.035 KILOCALORIES PER MOLE   Note 4

THE LAST 6 VIBRATIONS ARE THE TRANSLATION AND ROTATION MODES
THE FIRST THREE OF THESE BEING TRANSLATIONS IN X, Y, AND Z, RESPECTIVELY

NORMAL COORDINATE ANALYSIS

Note 5

ROOT NO.	1	2	3	4	5	6
	1209.89619	1214.68745	1490.52841	2114.52680	3255.93147	3302.10068
1	0.00000	0.00000	-0.02897	0.27081	0.00000	0.00047
2	-0.04861	0.00000	0.00000	0.00000	0.00287	0.00000
3	0.00000	-0.02732	0.00000	0.00000	0.00000	0.00000
4	0.00000	0.00000	-0.02661	-0.34468	0.00000	-0.04416
5	0.09729	0.00000	0.00000	0.00000	0.05937	0.00000
6	0.00000	0.10888	0.00000	0.00000	0.00000	0.00000
7	0.38049	0.00001	0.38850	-0.09565	-0.27942	0.25940
8	-0.19390	-0.00001	-0.26842	-0.16677	-0.37653	0.40112
9	0.00001	-0.43190	0.00000	0.00000	0.00000	0.00000
10	-0.38049	-0.00001	0.38850	-0.09565	0.27942	0.25940
11	-0.19390	0.00000	0.26842	0.16677	-0.37653	-0.40112
12	0.00001	-0.43190	0.00000	0.00000	0.00000	0.00000

ROOT NO.	7	8	9	10	11	12
	0.00047	-0.00052	-0.00066	21.05654	2.80744	3.83712
1	0.25000	0.00000	0.00000	0.00000	0.00000	0.00000
2	0.00000	-0.25000	0.00000	0.00000	0.00000	-0.14263
3	0.00000	0.00000	-0.25000	0.00000	-0.16682	0.00000
4	0.25000	0.00000	0.00000	0.00000	0.00000	0.00000
5	0.00000	-0.25000	0.00000	0.00000	0.00000	0.14214
6	0.00000	0.00000	-0.25000	0.00000	0.16625	0.00000
7	0.25000	0.00000	0.00000	0.00000	0.00000	-0.21588
8	0.00000	-0.25000	0.00000	0.00000	0.00000	0.28510
9	0.00000	0.00000	-0.25000	-0.50000	0.33346	0.00000
10	0.25000	0.00000	0.00000	0.00000	0.00000	0.21588
11	0.00000	-0.25000	0.00000	0.00000	0.00000	0.28510
12	0.00000	0.00000	-0.25000	0.50000	0.33346	0.00000

MASS-WEIGHTED COORDINATE ANALYSIS

Note 6

ROOT NO.	1	2	3	4	5	6
	1209.89619	1214.68745	1490.52841	2114.52680	3255.93147	3302.10068
1	0.00000	0.00000	-0.16876	0.66232	0.00000	0.00271
2	-0.26985	-0.00001	0.00000	0.00000	0.01649	0.00000
3	0.00000	-0.15005	0.00000	0.00000	0.00000	0.00000
4	0.00000	0.00000	-0.13434	-0.73040	0.00000	-0.22013
5	0.46798	0.00001	0.00000	0.00000	0.29524	0.00000
6	-0.00001	0.51814	0.00000	0.00000	0.00000	0.00000
7	0.53018	0.00002	0.56806	-0.05871	-0.40254	0.37456
8	-0.27018	-0.00001	-0.39249	-0.10237	-0.54244	0.57920
9	0.00002	-0.59541	0.00000	0.00000	0.00000	0.00000
10	-0.53018	-0.00001	0.56806	-0.05872	0.40254	0.37456
11	-0.27018	-0.00001	0.39249	0.10237	-0.54244	-0.57920
12	0.00002	-0.59541	0.00000	0.00000	0.00000	0.00000

ROOT NO.	7	8	9	10	11	12
	0.00040	-0.00044	-0.00062	21.05654	2.80744	3.83712
1	0.72996	0.00000	0.00000	0.00000	0.00000	0.00000
2	0.00000	-0.72996	0.00000	0.00000	0.00000	-0.62774
3	0.00000	0.00000	-0.72996	0.00000	-0.66681	0.00000
4	0.63247	0.00000	0.00000	0.00000	0.00000	0.00000
5	0.00000	-0.63247	0.00000	0.00000	0.00000	0.54204
6	0.00000	0.00000	-0.63247	0.00000	0.57578	0.00000

7	0.18321	0.00000	0.00000	0.00000	0.00000	-0.23848
8	0.00000	-0.18321	0.00000	0.00000	0.00000	0.31495
9	0.00000	0.00000	-0.18321	-0.70711	0.33455	0.00000
10	0.18321	0.00000	0.00000	0.00000	0.00000	0.23848
11	0.00000	-0.18321	0.00000	0.00000	0.00000	0.31495
12	0.00000	0.00000	-0.18321	0.70711	0.33455	0.00000

Note 7

DESCRIPTION OF VIBRATIONS

VIBRATION	1	ATOM PAIR	ENERGY CONTRIBUTION	RADIAL
FREQ.	1209.90	C 2 -- H 4	42.7% (79.4%)	12.6%
T-DIPOLE	0.7705	C 2 -- H 3	42.7%	12.6%
TRAVEL	0.1199	O 1 -- C 2	14.6%	0.0%
RED. MASS	1.9377			

VIBRATION	2	ATOM PAIR	ENERGY CONTRIBUTION	RADIAL
FRFQ.	1214.69	C 2 -- H 4	45.1% (62.3%)	0.0%
T-DIPOLE	0.0490	C 2 -- H 3	45.1%	0.0%
TRAVEL	0.1360	O 1 -- C 2	9.8%	0.0%
RED. MASS	1.5004			

VIBRATION	3	ATOM PAIR	ENERGY CONTRIBUTION	RADIAL
FREQ.	1490.53	C 2 -- H 3	49.6% (61.5%)	0.6%
T-DIPOLE	0.3443	C 2 -- H 4	49.6%	0.6%
TRAVEL	0.1846	O 1 -- C 2	0.9%	100.0%
RED. MASS	0.6639			

VIBRATION	4	ATOM PAIR	ENERGY CONTRIBUTION	RADIAL
FREQ.	2114.53	O 1 -- C 2	60.1% (100.5%)	100.0%
T-DIPOLE	2.9487	C 2 -- H 3	20.0%	17.7%
TRAVEL	0.0484	C 2 -- H 4	20.0%	17.7%
RED. MASS	6.7922			

VIBRATION	5	ATOM PAIR	ENERGY CONTRIBUTION	RADIAL
FREQ.	3255.93	C 2 -- H 3	49.5% (72.2%)	98.1%
T-DIPOLE	0.7518	C 2 -- H 4	49.5%	98.1%
TRAVEL	0.1174	O 1 -- C 2	1.0%	0.0%
RED. MASS	0.7508			

VIBRATION	6	ATOM PAIR	ENERGY CONTRIBUTION	RADIAL
FREQ.	3302.10	C 2 -- H 4	49.3% (69.8%)	95.5%
T-DIPOLE	0.3237	C 2 -- H 3	49.3%	95.5%
TRAVEL	0.1240	O 1 -- C 2	1.4%	100.0%
RED. MASS	0.6644			

SYSTEM IS A GROUND STATE

FORMALDEHYDE, MNDO ENERGY = -32.8819
DEMONSTRATION OF MOPAC - FORCE AND THERMODYNAMICS CALCULATION

MOLECULE IS NOT LINEAR

THERE ARE 6 GENUINE VIBRATIONS IN THIS SYSTEM
THIS THERMODYNAMICS CALCULATION IS LIMITED TO
MOLECULES WHICH HAVE NO INTERNAL ROTATIONS

CALCULATED THERMODYNAMIC PROPERTIES

TEMP. (K)	PARTITION FUNCTION	ENTHALPY CAL/MOL	HEAT CAPACITY CAL/K/MOL	ENTROPY CAL/K/MOL
298 VIB.	1.007	23.39443829	0.47838359	0.09150432
ROT.	709.	888.305	2.981	16.026
INT.	714.	911.700	3.459	16.117
TRA.	0.159E+27	1480.509	4.968	36.113
TOT.		2392.2084	8.4274	52.2300

TOTAL CPU TIME: 52.20 SECONDS

== MOPAC DONE ==

NOTE 1: All three words, ROT, FORCE, and THERMO are necessary in
order to obtain thermodynamic properties. In order to obtain results
for only one temperature, THERMO has the first and second arguments
identical. The symmetry number for the C2v point-group is 2.

NOTE 2: Internal coordinate derivatives are in Kcal/Angstrom or
Kcal/radian. Values of less than about 0.2 are quite acceptable.

NOTE 3: In larger calculations, the time estimates are useful. In
practice they are pessimistic, and only about 70% of the time estimated
will be used, usually. The principal moments of inertia can be directly
related to the microwave spectrum of the molecule. They are simple
functions of the geometry of the system, and are usually predicted with
very high accuracy.

NOTE 4: Zero point energy is already factored into the MNDO
parameterization. Force constant data are not printed by default. If
you want this output, specify LARGE in the keywords.

NOTE 5: Normal coordinate analysis has been extensively changed.
The first set of eigenvectors represent the 'normalized' motions of the
atoms. The sum of the speeds (not the velocities) of the atoms adds to
unity. This is verified by looking at the motion in the 'z' direction
of the atoms in vibration 2. Simple addition of these terms, unsigned,
adds to 1.0, whereas to get the same result for mode 1 the scalar of the
motion of each atom needs to be calculated first.

Users might be concerned about reproducibility. As can be seen from the vibrational frequencies from Version 3.00 to 5.00 given below, the main difference over earlier FORCE calculations is in the trivial frequencies.

Real Frequencies of Formaldehyde

Version 3.00	1209.96	1214.96	1490.60	2114.57	3255.36	3301.57
Version 3.10	1209.99	1215.04	1490.59	2114.57	3255.36	3301.58
Version 4.00	1209.88	1214.67	1490.52	2114.52	3255.92	3302.10
Version 5.00	1209.89	1214.69	1490.53	2114.53	3255.93	3302.10

Trivial Frequencies of Formaldehyde

	T(x)	T(y)	T(z)	R(x)	R(y)	R(z)
Version 3.00	-0.00517	-0.00054	-0.00285	57.31498	11.59518	9.01619
Version 3.10	-0.00557	0.00049	-0.00194	87.02506	11.18157	10.65295
Version 4.00	-0.00044	-0.00052	-0.00041	12.99014	-3.08110	-3.15427
Version 5.00	0.00040	-0.00044	-0.00062	21.05654	2.80744	3.83712

NOTE 6: Normal modes are not of much use in assigning relative importance to atoms in a mode. Thus in iodomethane it is not obvious from an examination of the normal modes which mode represents the C-I stretch. A more useful description is provided by the energy or mass-weighted coordinate analysis. Each set of three coefficients now represents the relative energy carried by an atom. (This is not strictly accurate as a definition, but is believed (by JJPS) to be more useful than the stricter definition.)

NOTE 7: Here there is a large change from Version 3.10. The following description of the coordinate analysis is given without rigorous justification. Again, the analysis, although difficult to understand, has been found to be more useful than previous descriptions.

On the left-hand side are printed the frequencies and transition dipoles. Underneath these are the reduced masses and idealized distances traveled which represent the simple harmonic motion of the vibration. The mass is assumed to be attached by a spring to an infinite mass. Its displacement is the travel.

The next column is a list of all pairs of atoms that contribute significantly to the energy of the mode. Across from each pair (next column) is the percentage energy contribution of the pair to the mode, calculated according to the formula described below.

FORMULA FOR ENERGY CONTRIBUTION

The total vibrational energy, T, carried by all pairs of bonded atoms in a molecule is first calculated. For any given pair of atoms, A and B, the relative contribution, R.C.(A,B), as a percentage, is given by the energy of the pair, P(A,B), times 100 divided by T, i.e.

$$R.C.(A,B) = 100P(A,B)/T$$

As an example, for formaldehyde the energy carried by the pair of atoms (C,O) is added to the energy of the two (C,H) pairs to give a total, T. Note that this total cannot be related to anything which is physically meaningful (there is obvious double-counting), but it is a convenient artifice. For mode 4, the C=O stretch, the relative contribution of the carbon-oxygen pair is 60.1%. It might be expected to be about 100% (after all, we envision the C=O bond as absorbing the photon); however, the fact that the carbon atom is vibrating implies that it is changing its position relative to the two hydrogen atoms. If the total vibrational energy, Ev (the actual energy of the absorbed photon, as distinct from T), were carried equally by the carbon and oxygen atoms, then the relative contributions to the mode would be C=O, 50% ; C-H, 25% ; C-H, 25%, respectively. This leads to the next entry, which is given in parentheses.

For the pair with the highest relative contribution (in mode 4, the C=O stretch), the energy of that pair divided by the total energy of the mode, Ev, is calculated as a percentage. This is the absolute contribution, A.C. as a percentage, to the total energy of the mode.

$$A.C.(A,B) \quad = \quad 100P(A,B)/Ev$$

Now the C=O is seen to contribute 100.5 percent of the energy. For this sort of partitioning only the sum of all A.C.'s must add to 100%, each pair can contribute more or less than 100%. In the case of a free rotator, e.g. ethane, the A.C. of any specific bonded pair to the total energy can be very high (several hundred percent).

It may be easier to view P/Ev as a contribution to the total energy of the mode, Ev. In this case the fact that P/Ev can be greater than unity can be explained by the fact that there are other relative motions within the molecule which make a negative contribution to Ev.

From the R.C.'s an idea can be obtained of where the energy of the mode is going; from the A.C. value the significance of the highest contribution can be inferred. Thus, in mode 4 all three bonds are excited, but because the C=O bond carries about 100% of the energy, it is clear that this is really a C=O bond stretch mode, and that the hydrogens are only going along for the ride.

In the last column the percentage radial motion is printed. This is useful in assigning the mode as stretching or bending. Any non-radial motion is de-facto tangential or bending.

To summarize: The new analysis is more difficult to understand, but is considered by the author (JJPS) to be the easiest way of describing what are often complicated vibrations.

NOTE 8:   In order, the thermodynamic quantities calculated are:

(1) The vibrational contribution,

(2) The rotational contribution,

(3) The sum of (1) and (2), this gives the internal contribution,

(4) The translational contribution.

For partition functions the various contributions are multiplied together.

5.3 EXAMPLE OF REACTION PATH WITH SYMMETRY

In this example, one methyl group in ethane is rotated relative to
the other and the geometry is optimized at each point. As the reaction
coordinate involves three Hydrogen atoms moving, symmetry is imposed to
ensure equivalence of all hydrogens.

```
Line  1:          SYMMETRY    T=600
Line  2:     ROTATION OF METHYL GROUP IN ETHANE
Line  3:     EXAMPLE OF A REACTION PATH CALCULATION
Line  4:     C
Line  5:     C    1.479146 1
Line  6:     H    1.109475 1  111.328433 1
Line  7:     H    1.109470 0  111.753160 0  120.000000  0   2  1  3
Line  8:     H    1.109843 0  110.103163 0  240.000000  0   2  1  3
Line  9:     H    1.082055 0  121.214083 0   60.000000 -1   1  2  3
Line 10:     H    1.081797 0  121.521232 0  180.000000  0   1  2  3
Line 11:     H    1.081797 0  121.521232 0  -60.000000  0   1  2  3
Line 12:     0    0.000000 0    0.000000 0    0.000000  0   0  0  0
Line 13:     3 1 4 5 6 7 8
Line 14:     3 2 4 5 6 7 8
Line 15:     6 7 7
Line 16:     6 11 8
Line 17:
Line 18:      70 80 90 100 110 120 130 140 150
```

Points to note:

(1) The dihedrals of the second and third hydrogens are not marked
for optimization: the dihedrals follow from point-group symmetry.

(2) All six C-H bond lengths and H-C-C angles are related by
symmetry: see lines 13 and 14.

(3) The dihedral on line 9 is the reaction coordinate, while the
dihedrals on lines 10 and 11 are related to it by symmetry functions on
lines 15 and 16 (see SYMMETRY for definitions of functions 1, 2, 7, and
11).

(4) Symmetry data are ended by a blank line.

(5) The reaction coordinate data are ended by the end of file.
Several lines of data are allowed.

(6) Whenever symmetry is used in addition to other data below the
geometry definition it will always follow the "blank line" immediately
following the geometry definition. The other data will always follow
the symmetry data.

CHAPTER 6

BACKGROUND

6.1 INTRODUCTION

While all the theory used in MOPAC is in the literature, so that in principle one could read and understand the algorithm, many parts of the code involve programming concepts or constructions which, while not of sufficient importance to warrant publication, are described here in order to facilitate understanding.

6.2 CORRECTION TO THE PEPTIDE LINKAGE

The residues in peptides are joined together by peptide linkages, -HNCO-. These linkages are almost flat, and normally adopt a trans configuration; the hydrogen and oxygen atoms being on opposite sides of the C-N bond. Experimentally, the barrier to interconversion in N-methyl acetamide is about 14 Kcal/mole, but all four methods within MOPAC predict a significantly lower barrier, PM3 giving the lowest value.

The low barrier can be traced to the tendency of semiempirical methods to give pyramidal nitrogens. The degree to which pyramidalization of the nitrogen atom is preferred can be seen in the following series of compounds.

Compound	MINDO/3	MNDO	AM1	PM3	Exp
Ammonia	Py	Py	Py	Py	Py
Aniline	Py	Py	Py	Py	Py
Formamide	Py	Py	Flat	Py	Py
Acetamide	Flat	Py	Flat	Py	Flat
N-methyl formamide	Flat	Py	Flat	Py	Flat
N-methyl acetamide	Flat	Flat	Flat	Py	Flat

To correct this, a molecular-mechanics correction has been applied. This consists of identifying the -R-HNCO- unit, and adding a torsion potential of form

$$Const*Sin(Theta)^2$$

where Theta is the X-N-C-O angle, X=R or H, and Const varies from method to method. This has two effects: there is a force constraining the nitrogen to be planar, and HNCO barrier in N-methyl acetamide is raised to 14.00 Kcal/mole. When the MM correction is in place, the nitrogen atom for all methods for the last three compounds shown above is planar. The correction should be user-transparent

Cautions

1. This correction will lead to errors of 0.5 - 1.5 Kcal/mole if the peptide linkage is made or broken in a reaction calculation.

2. If the correction is applied to formamide the nitrogen will be flat, contrary to experiment.

3. When calculating rotation barriers, take into account the rapid rehybridization which occurs. When the dihedral is 0 or 180 degrees the nitrogen will be planar (sp2), but at 90 degrees the nitrogen should be pyramidal, as the partial double bond is broken. At that geometry the true transition state involves motion of the nitrogen substituent so that the nitrogen in the transition state is more nearly sp2. In other words, a simple rotation of the HNCO dihedral will not yield the activation barrier, however it will be within 2 Kcal/mole of the correct answer. The 14 Kcal barrier mentioned earlier refers to the true transition state.

4. Any job involving a CONH group will require either the keyword NOMM or MMOK. If you do not want the correction to be applied, use the keyword "NOMM" (NO Molecular Mechanics).

6.3   LEVEL OF PRECISION WITHIN MOPAC

Several users have criticised the tolerances within MOPAC. The point made is that significantly different results have been obtained when different starting conditions have been used, even when the same conformer should have resulted. Of course, different results must be expected -- there will always be small differences -- nonetheless any differences should be small, e.g. heats of formation (H.o.F.) differences should be less than about 0.1 kcal/mole. MOPAC has been modified to allow users to specify a much higher precision than the default when circumstances warrant it.

Reasons for low precision

There are several reasons for obtaining low quality results. The most obvious cause of such errors is that for general work the default criteria will result in a difference in H.o.F. of less than 0.1 Kcal/mole. This is only true for fairly rigid systems, e.g. formaldehyde and benzene. For systems with low barriers to rotation or flat potential surfaces, e.g. aniline or water dimer, quite large H.o.F. errors can result.

Various Precision Levels

In normal (non-publication quality) work the default precision of MOPAC is recommended. This will allow reasonably precise results to be obtained in a reasonable time. Unless this precision proves unsatisfactory, use this default for all routine work.

The best way of controlling the precision of the geometry optimization and gradient minimization is by specifying a gradient norm which must be satisfied. This is done via the keyword GNORM=. Altering the GNORM automatically disables the other termination tests resulting in the gradient norm dominating the calculation. This works both ways: a GNORM of 20 will give a very crude optimization while a GNORM of 0.0001 will give a very precise optimization. The default GNORM is 1.0.

When the highest precision is needed, such as in exacting geometry work, or when you want results which cannot be improved, then use the combination keywords GNORM=0.0 and SCFCRT=1.D-NN; NN should be in the range 7-17. Increasing the SCF criterion (the default is SCFCRT=1.D-6) helps the line search routines by increasing the precision of the heat of formation calculation; however, it can lead to excessive run times, so take care. Also, there is an increased chance of not achieving an SCF when the SCF criterion is excessively increased.

Superficially, requesting a GNORM of zero might seem excessively stringent, but as soon as the run starts, it will be cut back to 0.0001. Even that might seem too stringent. The geometry optimization will continue to lower the energy, and hopefully the GNORM, but frequently it will not prove possible to lower the GNORM to 0.0001. If, after 10 cycles, the energy does not drop then the job will be stopped. At this point you have the best geometry that MOPAC, in its current form, can give.

If a slightly less than highest precision is needed, such as for normal publication quality work, set the GNORM to the limit wanted. For example, for a flexible system, a GNORM of 0.1 to 0.5 will normally be good enough for all but the most demanding work.

If higher than the default, but still not very high precision is wanted, then use the keyword PRECISE. This will tighten up various criteria so that higher than routine precision will be given.

If high precision is used, so that the printed GNORM is 0.000, and the resulting geometry resubmitted for one SCF and gradients calculation, then normally a GNORM higher than 0.000 will result. This is NOT an error in MOPAC: the geometry printed is only precise to six figures after the decimal point. Geometries need to be specified to more than six decimals in order to drive the GNORM to less than 0.000.

If you want to test MOPAC, or use it for teaching purposes, the GNORM lower limit of 0.0001 can be overridden by specifying LET, in which case you can specify any limit for GNORM. However, if it is too low the job may finish due to an irreducible minimum in the heat of formation being encountered. If this happens, the "STATIONARY POINT" message will be printed.

Finally there is a full analytical derivative function within MOPAC. These use STO-6G Gaussian wavefunctions because the derivatives of the overlap integral are easier to calculate in Gaussians than in STO's. Consequently, there will be a small difference in the calculated H.o.F.s when analytical derivatives are used. If there is any doubt about the accuracy of the finite derivatives, try using the analytical derivatives. They are a bit slower than finite derivatives but are more precise (a rough estimate is 12 figures for finite difference, 14 for analytical).

Some calculations, mainly open shell RHF or closed shell RHF with C.I. have untracked errors which prevent very high precision. For these systems GNORM should be in the range 1.0 to 0.1.

<div align="center">

How Large can a Gradient Be and
Still Be Acceptable?

</div>

A common source of confusion is the limit to which the GNORM should be reduced in order to obtain acceptable results. There is no easy answer, however a few guidelines can be given.

Firstly, reducing the GNORM to an arbitarily small number is not sensible. If we use a highly optimized geometry for formaldehyde - nine figure accuracy on bond lengths, something which cannot be done with a normal MOPAC - then a job run with 1SCF and GRADIENTS (see following data-files) will have no gradients higher that 0.000000.

Very Precise Formaldehyde Geometries

```
1SCF SCFCRT=1.D-13 GRADIENTS
Formaldehyde, for Demonstration Purposes
 Example of Geometry Optimized using Finite Difference Derivatives
O 0.0000000000 0   0.00000000 0   0.00000000 0   0   0   0
C 1.2164886796 1   0.00000000 0   0.00000000 0   1   0   0
H 1.1061255234 1 123.51907280 1   0.00000000 0   2   1   0
H 1.1061233045 1 123.50199135 1 179.99999999 0   2   1   3
0 0.00000000   0   0.000000   0   0.000000   0   0   0   0

ANALYT 1SCF SCFCRT=1.D-13 GRADIENTS
Formaldehyde, for Demonstration Purposes
 Example of Geometry Optimized using STO-6G Analytical Derivatives
O 0.0000000000 0   0.0000000 0   0.00000000 0   0   0   0
C 1.2164674433 1   0.0000000 0   0.00000000 0   1   0   0
H 1.1061618840 1 123.5197208 1   0.00000000 0   2   1   0
H 1.1061618840 1 123.5197208 1 180.00000000 0   2   1   3
0 0.0000000000 0   0.0000000 0   0.00000000 0   0   0   0
```

If any one numeral is changed, e.g. the C-O bond length is reduced to 1.216488679 or increased to 1.216488680 Angstroms, a change of less than 0.000000001 Angstroms, then finite gradients will be printed. Chemically, this change is utterly meaningless, and no significance should be attached to such numbers.

Note that in the finite-difference example the two C-H bonds and H-C-O angles are different. This is a real difference, and is due to as-yet untracked errors in MOPAC; however, the errors are quite small, about 6% of the difference between the two methods (finite difference and analytical derivatives).

As a guide, a GNORM of 0.1 is sufficient for all heat-of-formation work, and a GNORM of 0.01 for most geometry work. If the system is large, you may need to settle for a GNORM of 1.0 - 0.5.

This whole topic was raised by Dr. Donald B. Boyd of Lilly Research Laboratories, who provided unequivocal evidence for a failure of MOPAC and convinced me of the importance of increasing precision in certain circumstances.

6.4 CONVERGENCE TESTS IN SUBROUTINE ITER

Self-Consistency Test

The SCF iterations are stopped when two tests are satisfied. These are (1) when the difference in electronic energy, in eV, between any two consecutive iterations drops below the adjustable parameter, SELCON, and the difference between any three consecutive iterations drops below ten times SELCON, and (2) the difference in density matrix elements on two successive iterations falls below a preset limit, which is a multiple of SELCON.

SELCON is set initially to 0.00001 kcal/mole; this can be made 100 times smaller by specifying PRECISE or FORCE. It can be over-ridden by explicitly defining the SCF criterion via SCFCRT=1.D-12.

SELCON is further modified by the value of the gradient norm, if known. If GNORM is large, then a more lax SCF criterion is acceptable, and SCFCRT can be relaxed up to 50 times its default value. As the gradient norm drops, the SCF criterion returns to its default value.

The SCF test is performed using the energy calculated from the Fock matrix which arises from a density matrix, and not from the density matrix which arises from a Fock. In the limit, the two energies would be identical, but the first converges faster than the second, without loss of precisior

6.5   CONVERGENCE IN SCF CALCULATION

A brief description of the convergence techniques used in subroutine ITER follows.

ITER, the SCF calculation, employs six methods to achieve a self-consistent field. In order of usage, these are:

(a) Intrinsic convergence by virtue of the way the calculation is carried out. Thus a trial Fock gives rise to a trial density matrix, which in turn is used to generate a better Fock matrix.

This is normally convergent, but many exceptions are known. The main situations when the intrinsic convergence does not work are:

> (1) A bad starting density matrix. This normally occurs when the default starting density matrix is used. This is a very crude approximation, and is only used to get the calculation started. A large charge is generated on an atom in the first iteration, the second iteration overcompensates, and an oscillation is generated.

> (2) The equations are only very slowly convergent. This can be due to a long-lived oscillation or to a slow transfer of charge.

(b) Oscillation damping. If, on any two consecutive iterations, a density matrix element changes by more than 0.05, then the density matrix element is set equal to the old element shifted by 0.05 in the direction of the calculated element. Thus, if on iterations 3 and 4 a certain density matrix element was 0.55 and 0.78, respectively, then the element would be set to 0.60 (=0.55+0.05) on iteration 4. The density matrix from iteration 4 would then be used in the construction of the next Fock matrix. The arrays which hold the old density matrices are not filled until after iteration 2. For this reason they are not used in the damping before iteration 3.

(c) Three-point interpolation of the density matrix. Subroutine CNVG monitors the number of iterations, and if this is exactly divisible by three, and certain other conditions relating to the density matrices are satisfied, a three-point interpolation is performed. This is the default converger, and is very effective with normally convergent calculations. It fails in certain systems, usually those where significant charge build-up is present.

(d) Energy-level shift technique. The virtual M.O. energy levels are normally shifted to more positive energy. This has the effect of damping oscillations, and intrinsically divergent equations can often be changed to intrinsically convergent form. With slowly-convergent systems the virtual M.O. energy levels can be moved to a more negative value.

The precise value of the shift used depends on the behavior of the iteration energy. If it is dropping, then the HOMO-LUMO gap is reduced, if the iteration energy rises, the gap is increased rapidly.

(e) Pulay's method. If requested, when the largest change in density matrix elements on two consecutive iterations has dropped below 0.1, then routine CNVG is abandoned in favor of a multi-Fock matrix interpolation. This relies on the fact that the eigenvectors of the density and Fock matrices are identical at self-consistency, so [P.F]=0 at SCF. The extent to which this condition does not occur is a measure of the deviance from self-consistency. Pulay's method uses this relationship to calculate that linear combination of Fock matrices which minimize [P.F]. This new Fock matrix is then used in the SCF calculation.

Under certain circumstances, Pulay's method can cause very slow convergence, but sometimes it is the only way to achieve a self-consistent field. At other times the procedure gives a ten-fold increase in speed, so care must be exercised in its use. (invoked by the keyword PULAY)

(f) The Camp-King converger. If all else fails, the Camp-King converger is just about guaranteed to work every time. However, it is time-consuming, and therefore is invoked as a last resort.

It evaluates that linear combination of old and current eigenvectors which minimize the total energy. One of its strengths is that systems which otherwise oscillate due to charge surges, e.g. CHO-H, the C-H distance being very large, will converge using this very sophisticated converger.

6.6   CAUSES OF FAILURE TO ACHIEVE AN SCF

In a system where a biradical can form, such as ethane decomposing into two CH_3 units, the normal RHF procedure can fail to go self-consistent. If the system has marked biradicaloid character, then BIRADICAL or UHF and TRIPLET can often prove successful. These options rely on the assumption that two unpaired electrons can represent the

open shell part of the wave-function.

Consider H-Cl, with the interatomic distance being steadily increased. At first the covalent bond will be strong, and a self-consistent field is readily obtained. Gradually the bond will become more ionic, and eventually the charge on chlorine will become very large. The hydrogen, meanwhile, will become very electropositive, and there will be an increased energy advantage to any one electron to transfer from chlorine to hydrogen. If this in fact occurred, the hydrogen would suddenly become very electron-rich and would, on the next iteration, lose its extra electron to the chlorine. A sustained oscillation would then be initiated. To prevent this, if BIRADICAL is specified, exactly one electron will end up on hydrogen. A similar result can be obtained by specifying TRIPLET in a UHF calculation.

6.7 TORSION OR DIHEDRAL ANGLE COHERENCY

MOPAC calculations do not distinguish between enantiomers, consequently the sign of the dihedrals can be multiplied by -1 and the calculations will be unaffected. However, if chirality is important, a user should be aware of the sign convention used.

The dihedral angle convention used in MOPAC is that defined by Klyne and Prelog in Experientia 16, 521 (1960). In this convention, four atoms, AXYB, with a dihedral angle of 90 degrees, will have atom B rotated by 90 degrees clockwise relative to A when X and Y are lined up in the direction of sight, X being nearer to the eye. In their words, "To distinguish between enantiomeric types the angle 'tau' is considered as positive when it is measured clockwise from the front substituent A to the rear substituent B, and negative when it is measured anticlockwise." The alternative convention was used in all earlier programs, including QCPE 353.

6.8 VIBRATIONAL ANALYSIS

Analyzing normal coordinates is very tedious. Users are normally familiar with the internal coordinates of the system they are studying, but not familiar with the cartesian coordinates. To help characterize the normal coordinates, a very simple analysis is done automatically, and users are strongly encouraged to use this analysis first, and then to look at the normal coordinate eigenvectors.

In the analysis, each pair of bonded atoms is examined to see if there is a large relative motion between them. By bonded is meant within the Van der Waals' distance. If there is such a motion, the indices of the atoms, the relative distance in Angstroms, and the percentage radial motion are printed. Radial plus tangential motion adds to 100%, but as there are two orthogonal tangential motions and only one radial, the radial component is printed.

6.9   A NOTE ON THERMOCHEMISTRY

By

Tsuneo Hirano
Department of Synthetic Chemistry
Faculty of Engineering
University of Tokyo
Hongo, Bunkyo-ku, Tokyo, Japan

1) Basic Physical Constants
"Quantities, Units and Symbols in Physical Chemistry,"
Blackwell Scientific Publications Ltd, Oxford OX2 0EL, UK, 1987
(IUPAC, based on CODATA of ICSU, 1986).   pp 81-82.

Speed of light c = 2.997 92458 D10   cm/s (Definition)

Boltzmann constant k = R/Na

                              = 1.380 658 D-23   J/K

                              = 1.380 658 D-16   erg/K

Planck constant h = 6.626 0755 D-34   J s

                              = 6.626 0755 D-27   erg s

Gas constant R = 8.314 510   J/mol/K

                              = 1.987 216   cal/mol/K

Avogadoro number   Na = 6.022 1367 D23   /mol

Volume of 1 mol of gas   V0 = 22.414 10 1/mol (at 1 atm, 25 C)

1 J = 1.D7 erg

1 kcal = 4.184 kJ   (Definition)

1 eV = 23.060 6 kcal/mol

1 a.u. = 27.211 35 eV/mol = 627.509 6 kcal/mol

1 cm-1 = 2.859 144 cal/mol (= Na h c / 4.184D7)

1 atm = 1.013 25 D5 Pa = 1.013 25 D6 dyn/cm**2 (Definition)

* * * * * *

Moment of inertia: I

1 amu angstrom**2 = 1.660 540 D-40   g cm**2

Rotational constants: A, B, and C (e.g. A = h/(8*pai*pai*I))

$$A(\text{in MHz}) = 5.053\ 791\ D5\ /\ I(\text{in amu angstrom}**2)$$

$$A(\text{in cm-1}) = 5.053\ 791\ D5/\ c/\ I(\text{in amu angstrom}**2)$$

$$= 16.857\ 63\ /\ I(\text{in amu angstrom}**2)$$

2) Thermochemistry from ab initio MO methods.

Ab initio MO methods provide total energies, Eeq, as the sum of electronic and nuclear-nuclear repulsion energies for molecules, isolated in vacuum, without vibration at 0 K.

$$Eeq = Eel + Enuclear-nuclear \tag{1}$$

From the 0 K-potential surface and usin the harmonic oscillator approximation, we can calculate the vibrational frequencies, vi, of the normal modes of vibration. Using these, we can calculate vibrational, rotational and translational contributions to the thermodynamic quantities such as the partition function and heat capacity which arise from heating the system from 0 to T K.

    Q: partition function
    E: energy
    S: entropy
    C: heat capacity

[Vibration]

$$Qvib = \text{pai over i } \{\ 1/(1 - \exp(-hvi/kT))\} \tag{2}$$

Evib, for a molecule at the temperature T as

$$Evib = \text{sum over i } \{(1/2)hvi + hvi*\exp(-hvi/kT)/(1 - \exp(-hvi/kT))\}$$

$$\tag{3}$$

where h is the Planck constant, vi the i-th normal vibration frequency, and k the Boltzmann constant. For 1 mole of molecules, Evib should be multiplied by the Avogadro number Na(= gas constant R/k). Thus,
$$Evib = Na * \text{sum over i } \{(1/2)hvi$$

$$+ hvi*\exp(-hvi/kT)/(1-\exp(-hvi/kT))\} \tag{4}$$

Note that the first term in Eq. 4 is the Zero-point vibration energy. Hence, the second term in Eq. 4 is the additional vibrational contribution due to the temperature increase from 0 K to T k. Namely,

$$Evib = Ezero + Evib(0-->T) \tag{5}$$

$$Ezero = Na * sum\ over\ i\ \{(1/2)hvi\}, \tag{6}$$

$$Evib(0-->T) = Na * sum\ over\ i\ \{hvi*exp(-hvi/kT)/(1 - exp(-hvi/kT))\}. \tag{7}$$

The value of Evib from GAUSSIAN 82 and 86 includes Ezero as defined by Eqs. 4 - 7.

$$Svib = R\ sum\ ovrer\ i\ \{(hvi/kT)*exp(-hvi/kT)/(1 - exp(-hvi/kT))$$
$$- ln(1 - exp(-hvi/kT))\} \tag{8}$$

$$Cviv = R\ sum\ over\ i\ \{((hvi/kT)**2)\ exp(-hvi/kT)/$$
$$(1 - exp(-hvi/kT))**2\} \tag{9}$$

At temperature T (>0 K), a molecule rotates about the x, y, and z-axes and translates in x, y, and z-directions. By assuming the equipartition of energy, energies for rotation and translation, Erot and Etr, are calculated.
[Rotation]

(sym) is symmetry number.
I is moment of inertia.
IA, IB, and IC are moments of inertia about A, B, and C axes.

<linear molecule>

$$Qrot = (1/(sym))[8(pai**2)\ I\ kT/\ h**2] \tag{10}$$

$$Erot = (2/2)RT \tag{11}$$

$$Srot = R\ ln\ [(1/(sym))*(8(pai**2)\ I/h**2)*kT] + R \tag{12}$$

$$= R\ ln\ I + R\ ln\ T - R\ ln(sym) - 4.349\ 203 \tag{13}$$

where

$$-4.349\ 203 = R\ ln\{8(pai**2)(1/Na)(1.D-8)**2\ k/h**2\} + R.$$

$$Crot = (2/2)R \tag{14}$$

<non-linear molecule>

$$Qrot = (pai**0.5/(sym))\ [8(pai**2)kT/h**2]**(3/2)\ (IA\ IB\ IC)**(1/2)$$

$$= (pai**0.5/(sym))\ \{[8(pai**2)(IA)c/h]*[8(pai**2)(IB)c/h]*$$

$$[8(pai**2)(IC)c/h]\}**(1/2)\ (kT/hc)**(3/2) \tag{15}$$

$$Erot = (3/2)RT \tag{16}$$

$$Srot = (R/2) \ln \{(pai/(sym)**2) [8(pai**2)(IA)c/h]*$$

$$[8(pai**2)(IB)c/h]*[8(pai**2)(IC)c/h]*[(kT/hc)**3]\}$$

$$+ (3/2)R \tag{17}$$

$$= (R/2) \ln (IA \ IB \ IC) + (3/2) \ R \ln T - R \ln (sym) - 5.386 \ 3921$$

Here, -5.386 3921 is calculated as

$$R \ln \{(((1.D-8)**2/Na)**3)**(1/2) (2**9 \ pai**7 \ k*3)**(1/2) / h**3\} + (3/2)R.$$

$$Crot = (3/2)R \tag{18}$$

[Translation]

M is Molecular weight.

$$Qtra = \{[2 \ (pai) \ (M/Na) \ kT]**(1/2) / h\}**3 \tag{19}$$

$$Etra = (3/2)RT \tag{20}$$

(or Htra = (5/2)RT due to the pV term (cf. H = U + pV))

$$Stra = R \{ (5/2) + (3/2)\ln [2(pai)k/h**2] + \ln k + (3/2)\ln (M/Na)$$

$$+ (5/2)\ln T - \ln p \} \tag{21}$$

$$= (5/2)R \ln T + (3/2)R \ln M - R \ln p - 2.31482 \tag{22}$$

$$Ctra = (5/2)R \tag{23}$$

The internal energy U at T is

$$U = Eeq + [Evib + Erot + Etra] \tag{24}$$

or

$$U = Eeq + [(Ezero + Evib(0-->T)) + Erot + Etra] \tag{25}$$

Enthalpy H for one mole of gas is defined as

$$H = U + pV \tag{26}$$

Assumption of an ideal gas (i.e. pV = RT) leads to

$$H = U + pV = U + RT \tag{27}$$

Thus, Gibbs free energy G can be calculated as

$$G = H - T*S(0-->T) \tag{28}$$

2)Thermochemistry in MOPAC

It should be noted that MO parameters for MINDO/3, MNDO, AM1 and PM3 are optimized so as to reproduce the experimental heat of formation (i.e. standard enthalpy of formation or the enthalpy change to form a mole of compound at 25 degrees C from its elements in their standard state) as well as observed geometries (mostly at 25 degrees C), and not to reproduce the Eeq and equilibrium geometry at 0 K.

In this sense, Escf (defined as Heat of formation), force constants, normal vibration frequencies etc are all related to the values at 25 degree C, not to 0 K!!!!! Therefore, the Ezero calculated in FORCE is not the true Ezero. Its use as Ezero should be made at your own risk, bearing in mind the situation discussed above.

Since Escf is standard enthalpy of formation (at 25 degree C),

$$Escf = [Eeq + Ezero + Evib(0-->298.15) + Erot + Etra + pV]$$

$$+ sum [- Electronic\ energy\ of\ atom + Delta\_H\ of\ formation\ of\ atom].$$

$$\tag{29}$$

To avoid the complication arising from the definition of Escf, within the thermodynamics calculation the Standard Enthalpy of Formation, DeltaH, is calculated by

$$Delta\_H = Escf + (HT - H298). \tag{30}$$

Here, Escf is the heat of formation (at 25 degree C) given in the output list, and HT and H298 are the enthalpy contributions for the increase of the temperature from 0 K to T and 298.15, respectively. In other words, the enthalpy of formation is corrected for the difference in temperature from 298.15 K to T. The method of calculation for T and H298 will be given below.

In MOPAC, the variables defined below are used

$$C1 = hc / kT \tag{31}$$

$$Wi\ (in\ cm-1)\ (i.e.\quad vi = Wi*c\ ), \tag{32}$$

$$EWJ = exp(\ -hvi/kT) = exp(\ -Wi*hc/kT) = exp(- Wi*C1) \tag{33}$$

$$A,\ B,\ and\ C\ in\ cm-1\ (i.e.\quad A = [h/(8(pai**2)IA*c)]), \tag{34}$$

Energy and Enthalpy in cal/mol, and Entropy in cal/mol/K.

Thus, eqs. 2 - 28 can be written as follows.

[Vibration]

$$Qviv = pai\ over\ i\ \{1\ /\ (1\ -\ EWJ)\} \tag{35}$$

$$E0 = [0.5\ Na\ h\ c/(4.184\ D7)]\ sum\ over\ i\ \{Wi\} \tag{36}$$

$$= 1.429\ 572\ *\ sum\ over\ i\ \{Wi\} \tag{37}$$

$$Evib(0->T) = Na\ h\ c\ sum\ over\ i\ \{\ Wi*EWJ\ /\ (1\ -\ EWJ)\ \}$$

$$= (R/k)\ h\ c\ sum\ over\ i\ \{\ Wi*EWJ\ /\ (1\ -\ EWJ)\ \} \tag{38}$$

$$Svib = R\ (hc/kT)\ sum\ over\ i\ \{WI*EWJ/(1\ -\ EWJ)\}$$

$$-\ R\ sum\ over\ i\ \{ln\ (1\ -\ EWJ)\}$$

$$= R\ C1\ sum\ over\ i\ \{WI*EWJ/(1\ -\ EWJ)\}$$

$$-\ R\ sum\ over\ i\ \{ln\ (1\ -\ EWJ)\} \tag{39}$$

$$Cvib = R\ (hc/kT)**2\ sum\ over\ i\ \{\ Wi**2\ EWJ\ /(1-\ EWJ)**2\ \}$$

$$= R\ C1**2\ sum\ over\ i\ \{\ Wi**2\ EWJ\ /(1-\ EWJ)**2\ \} \tag{40}$$

[Rotation]

<linear molecule>

$$Qrot = (1/(sym))\ (1/A)\ (kT/hc)$$

$$= 1/[(sym)\ A\ C1] \tag{41}$$

$$Erot = (2/2)RT \tag{42}$$

$$Srot = R\ ln\ [(1/(sym))\ (1/A)\ (kT/hc)]\ +\ R$$

$$= R\ ln\ [1/(\ (sym)*A*C1\ )]\ +\ R$$

$$= R\ ln\ [\ Tk/(\ h*c*A*sym)]\ +\ R \tag{43}$$

$$Crot = (2/2)R \tag{44}$$

<non-linear molecule>

$$Qrot = [\ pai\ /\ (A\ B\ C\ C1**3)]**(1/2)\ /\ (sym) \tag{45}$$

$$Erot = (3/2)RT \tag{46}$$

$$Srot = (R/2) \ln \{ (1/(sym)**2) (1/A) (1/B) (1/C) (pai) (kT/hc)**3 \}$$

$$+ (3/2)R$$

$$= 0.5R \{ 3 \ln (kT/hc) - 2 \ln (sym) + \ln (pai/(A \; B \; C)) + 3\}$$

(47)

$$= 0.5R \{ -3 \ln Cl -2 \ln (sym) + \ln (pai/(A \; B \; C)) + 3\}$$

$$Crot = (3/2)R \qquad\qquad (48)$$

[Translation]

$$Qtra = [(2 \; pai \; (M/Na) \; kT)**(1/2) \; / \; h]**3$$

$$= [(2 \; pai \; M \; k \; T * 1.660540D-24)**(1/2) \; /h]**3 \qquad (49)$$

$$Etra = (3/2)RT \qquad\qquad (50)$$

$$Htra = (3/2)RT + pV = (5/2)RT \quad (cf. \; pV = RT) \qquad (51)$$

$$Stra = (R/2) [5 \ln T + 3 \ln M] - 2.31482 \quad (cf. \; p = 1 \; atm)$$

$$= 0.993608 [5 \ln T + 3 \ln M] - 2.31482 \qquad (52)$$

In MOPAC,

$$Hvib = Evib(0-->T) \qquad\qquad (53)$$

(Note: Ezero is <not> included in Hvib.
Wi is not derived from force-constants at 0 K)

and

$$HT \quad = [Hvib + Hrot + Htra] \qquad for \; T. \qquad (54)$$

$$H298 = [Hvib + Hrot + Htra] \qquad for \; T = 298.15. \qquad (55)$$

Note that HT (and H298) is equivalent to

$$[(Evib - Ezero) + Erot + (Etra + pV)] \qquad (56)$$

except that the normal frequencies are those obtained from force constants at 25 degree C, or at least not at 0 K.

Thus, Standard Enthalpy of Formation, DeltaH, can be calculated according to Eqs. 25, 26 and 29, as shown in Eq. 30;

$$Delta\_H = Escf + (HT - H298) \qquad (57)$$

Note that Ezero is already counted in Escf (see Eq. 29).

By using Eq. 27, Standard Internal Energy of Formation, DeltaU, can be calculated as

$$Delta\_U = Delta\_H - R(T - 298.15). \tag{58}$$

Standard Gibbs Free-Energy of Formation, DeltaG, can be calculated by taking the difference from that for the isomer or that at different temperature,

$$Delta\_G = [Delta\_H - T*S] \text{ for the state under consideration}$$

$$- [Delta\_H - T*S] \text{ for reference state.} \tag{59}$$

Taking the difference is necessary to cancel the unknown values of standard entropy of formation for the constituent elements.

6.10   REACTION COORDINATES

The Intrinsic Reaction Coordinate method pioneered and developed by Mark Gordon has been incorporated in a modified form into MOPAC. As this facility is quite complicated all the keywords associated with the IRC have been grouped together in this section.

Definitions of Terms

DRC

The Dynamic Reaction Coordinate is the path followed by all the atoms in a system assuming conservation of energy, i.e., as the potential energy changes the kinetic energy of the system changes in exactly the opposite way so that the total energy (kinetic plus potential) is a constant. If started at a ground state geometry, no significant motion should be seen. Similarly, starting at a transition state geometry should not produce any motion - after all it is a stationary point and during the lifetime of a calculation it is unlikely to accumulate enough momentum to travel far from the starting position.

In order to calculate the DRC path from a transition state, either an initial deflection is necessary or some initial momentum must be supplied.

Because of the time-dependent nature of the DRC the time elapsed since the start of the reaction is meaningful, and is printed.

IRC

The Intrinsic Reaction Coordinate is the path followed by all the atoms in a system assuming all kinetic energy is completely lost at every point, i.e., as the potential energy changes the kinetic energy generated is annihilated so that the total energy (kinetic plus

potential) is always equal to the potential energy only.

The IRC is intended for use starting with the transition state geometry. A normal coordinate is chosen, usually the reaction coordinate, and the system is displaced in either the positive or negative direction along this coordinate. The internal modes are obtained by calculating the mass-weighted Hessian matrix in a force calculation and translating the resulting cartesian normal mode eigenvectors to conserve momentum. That is, the initial cartesian coordinates are displaced by a small amount proportional to the eigenvector coefficients plus a translational constant; the constant is required to ensure that the total translational momentum of the system is conserved as zero. At the present time there may be small residual rotational components which are not annihilated; these are considered unimportant.

General Description of the DRC and IRC.

As the IRC usually requires a normal coordinate, a force constant calculation normally has to be done first. If IRC is specified on its own a normal coordinate is not used and the IRC calculation is performed on the supplied geometry.

A recommended sequence of operations to start an IRC calculation is as follows:

1. Calculate the transition state geometry. If the T/S is not first optimized, then the IRC calculation may give very misleading results. For example, if NH3 inversion is defined as the planar system but without the N-H bond length being optimized the first normal coordinate might be for N-H stretch rather than inversion. In that case the IRC will relax the geometry to the optimized planar structure.

2. Do a normal FORCE calculation, specifying ISOTOPE in order to save the FORCE matrices. Do not attempt to run the IRC directly unless you have confidence that the FORCE calculation will work as expected. If the IRC calculation is run directly, specify ISOTOPE anyway: that will save the FORCE matrix and if the calculation has to be re-done then RESTART will work correctly.

3. Using IRC=n and RESTART run the IRC calculation. If RESTART is specified with IRC=n then the restart is assumed to be from the FORCE calculation. If RESTART is specified without IRC=n, say with IRC on its own, then the restart is assumed to be from an earlier IRC calculation that was shut down before going to completion.

A DRC calculation is simpler in that a force calculation is not a prerequisite; however, most calculations of interest normally involve use of an internal coordinate. For this reason IRC=n can be combined with DRC to give a calculation in which the initial motion (0.3Kcal worth of kinetic energy) is supplied by the IRC, and all subsequent motion obeys conservation of energy. The DRC motion can be modified in three ways:

1. It is possible to calculate the reaction path followed by a system in which the generated kinetic energy decays with a finite half-life. This can be defined by DRC=n.nnn, where n.nnn is the half-life in femtoseconds. If n.nn is 0.0 this corresponds to infinite damping simulating the IRC. A limitation of the program is that time only has meaning when DRC is specified without a half-life.

2. Excess kinetic energy can be added to the calculation by use of KINETIC=n.nn. After the kinetic energy has built up to 0.2kcal/mole or if IRC=n is used then n.nn Kcal/mole of kinetic energy is added to the system. The excess kinetic energy appears as a velocity vector in the same direction as the initial motion.

3. The RESTART file <filename>.RES can be edited to allow the user to modify the velocity vector or starting geometry. This file is formatted.

Frequently DRC leads to a periodic, repeating orbit. One special type - the orbit in which the direction of motion is reversed so that the system retraces its own path - is sensed for and if detected the calculation is stopped after exactly one cycle. If the calculation is to be continued, the keyword GEO-OK will allow this check to be by-passed.

Due to the potentially very large output files that the DRC can generate extra keywords are provided to allow selected points to be printed. After the system has changed by a preset amount the following keywords can be used to invoke a print of the geometry.

KeyWord	Default	User Specification
X-PRIO	0.05 Angstroms	X-PRIORITY=n.nn
T-PRIO	0.10 Femtoseconds	T-PRIORITY=n.nn
H-PRIO	0.10 Kcal/mole	H-PRIORITY=n.nn

Option to allow only extrema to be output

In the geometry specification, if an internal coordinate is marked for optimization then when that internal coordinate passes through an extremum a message will be printed and the geometry output.

Difficulties can arise from the way internal coordinates are processed. The internal coordinates are generated from the cartesian coordinates, so an internal coordinate supplied may have an entirely different meaning on output. In particular the connectivity may have changed. For obvious reasons dummy atoms should not be used in the supplied geometry specification. If there is any doubt about the internal coordinates or if the starting geometry contains dummy atoms then run a 1SCF calculation specifying XYZ. This will produce an ARC file with the "ideal" numbering - the internal numbering system used by MOPAC. Use this ARC file to construct a data file suitable for the DRC or IRC.

Notes

1. Any coordinates marked for optimization will result in only extrema being printed.

2. If extrema are being printed then kinetic energy extrema will also be printed.

Keywords for use with the IRC and DRC

1. Setting up the transition state: NLLSQ SIGMA.

2. Constructing the FORCE matrix: FORCE or IRC=n, ISOTOPE, LET.

3. Starting an IRC: RESTART and IRC=n, T-PRIO, X-PRIO, H-PRIO.

4. Starting a DRC: DRC or DRC=n.nn, KINETIC=n.nn.

5. Starting a DRC from a transition state: (DRC or DRC=n) and IRC=n, KINETIC=n.

6. Restarting an IRC: RESTART and IRC.

7. Restarting a DRC: RESTART and (DRC or DRC=n.nn).

8. Restarting a DRC starting from a transition state: RESTART and (DRC or DRC=n.nn).

Other keywords, such as T=nnn or GEO-OK can be used anytime.

Examples of DRC/IRC data

Use of the IRC/DRC facility is quite complicated. In the following examples various "reasonable" options are illustrated for a calculation on water.

It is assumed that an optimized transition-state geometry is available.

Example 1: A Dynamic Reaction Coordinate, starting at the transition state for water inverting, initial motion opposite to the transition normal mode, with 6kcal of excess kinetic energy added in. Every point calculated is to be printed (Note all coordinates are marked with a zero, and T-PRIO, H-PRIO and X-PRIO are all absent). The results of an earlier calculation using the same keywords is assumed to exist. The earlier calculation would have constructed the force matrix. While the total cpu time is specified, it is in fact redundant in that the calculation will run to completion in less than 600 seconds.


```
KINETIC=6 RESTART   IRC=-1 DRC T=600
    WATER

 H    0.000000  0      0.000000  0    0.000000  0   0  0  0
 O    0.911574  0      0.000000  0    0.000000  0   1  0  0
 H    0.911574  0    180.000000  0    0.000000  0   2  1  0
 O    0.000000  0      0.000000  0    0.000000  0   0  0  0
```

Example 2: An Intrinsic Reaction Coordinate calculation. Here the restart is from a previous IRC calculation which was stopped before the minimum was reached. Recall that RESTART with IRC=n implies a restart from the FORCE calculation. Since this is a restart from within an IRC calculation the keyword IRC=n has been replaced by IRC. IRC on its own (without the "=n") implies an IRC calculation from the starting position - here the RESTART position - without initial displacement.


```
    RESTART   IRC   T=600
    WATER

 H    0.000000  0      0.000000  0    0.000000  0   0  0  0
 O    0.911574  0      0.000000  0    0.000000  0   1  0  0
 H    0.911574  0    180.000000  0    0.000000  0   2  1  0
 O    0.000000  0      0.000000  0    0.000000  0   0  0  0
```

Output Format for IRC and DRC

The IRC and DRC can produce several different forms of output. Because of the large size of these outputs, users are recommended to use search functions to extract information. To facilitate this, specific lines have specific characters. Thus, a search for the "%" symbol will

summarize the energy profile while a search for "AA" will yield the coordinates of atom 1, whenever it is printed. The main flags to use in searches are:

SEARCH FOR	YIELDS
'% '	Energies for all points calculated, excluding extrema
'%M'	Energies for all turning points
'%MAX'	Energies for all maxima
'%MIN'	Energies for all minima
'%'	Energies for all points calculated
'AA*'	Internal coordinates for atom 1 for every point
'AE*'	Internal coordinates for atom 5 for every point
'123AB*'	Internal coordinates for atom 5 for point 123

As the keywords for the IRC/DRC are interdependent, the following list of keywords illustrates various options.

Keyword	Resulting Action
DRC	The Dynamic Reaction Coordinate is calculated. Energy is conserved, and no initial impetus.
DRC=0.5	In the DRC kinetic energy is lost with a half-life of 0.5 femtoseconds.
DRC=-1.0	Energy is put into a DRC with an half-life of -1.0 femtoseconds, i.e. the system gains energy.
IRC	The Intrinsic Reaction Coordinate is calculated. No initial impetus is given. Energy not conserved.
IRC=-4	The IRC is run starting with an impetus in the negative of the 4th normal mode direction. The impetus is one quantum of vibrational energy.
IRC=1 KINETIC=1	The first normal mode is used in an IRC, with the initial impetus being 1.0Kcal/mole.
DRC KINETIC=5	In a DRC, after the velocity is defined, 5 Kcal of kinetic energy is added in the direction of the initial velocity.
IRC=1 DRC KINETIC=4	After starting with a 4 kcal impetus in the direction of the first normal mode, energy is conserved.

Instead of every point being printed, the option exists to print specific points determined by the keywords T-PRIORITY, X-PRIORITY and H-PRIORITY. If any one of these words is specified, then the calculated points are used to define quadratics in time for all variables normally printed. In addition, if the flag for the first atom is set to T then all kinetic energy turning points are printed. If the flag for any other internal coordinate is set to T then, when that coordinate passes through an extremum, that point will be printed. As with the PRIORITYs, the point will be calculated via a quadratic to minimize non-linear errors.

N.b.: Quadratics are unstable in the regions of inflection points, in these circumstances linear interpolation will be used. A result of this is that points printed in the region of an inflection may not correspond exactly to those requested. This is not an error and should not affect the quality of the results.

Test of DRC - Verification of Trajectory Path

Introduction: Unlike a single-geometry calculation or even a geometry optimization, verification of a DRC trajectory is not a simple task. In this section a rigorous proof of the DRC trajectory is presented; it can be used both as a test of the DRC algorithm and as a teaching exercise. Users of the DRC are asked to follow through this proof in order to convince themselves that the DRC works as it should.

Part 1:   The Nitrogen Molecule

For the nitrogen molecule and using MNDO, the equilibrium distance is 1.103682 Angstroms, the heat of formation is 8.79666 Kcal/mole and the vibrational frequency is 2740.4 cm(-1). For small displacements, the energy curve versus distance is parabolic and the gradient curve is approximately linear, as is shown in the following table. A nitrogen molecule is thus a good approximation to a harmonic oscillator.

STRETCHING CURVE FOR NITROGEN MOLECULE

N-N DIST (Angstroms)	H.O.F. (Kcal/mole)	GRADIENT (Kcal/mole/Angstrom)
1.1180	9.242203	61.440366
1.1170	9.182830	57.302349
1.1160	9.127607	53.141739
1.1150	9.076555	48.958460
1.1140	9.029698	44.752436
1.1130	8.987058	40.523588
1.1120	8.948659	36.271841
1.1110	8.914523	31.997118
1.1100	8.884673	27.699340
1.1090	8.859133	23.378431
1.1080	8.837925	19.034313
1.1070	8.821073	14.666909
1.1060	8.808599	10.276141
1.1050	8.800529	5.861931
1.1040	8.796884	1.424201
1.1030	8.797689	-3.037127
1.1020	8.802967	-7.522131
1.1010	8.812742	-12.030889
1.1000	8.827037	-16.563480
1.0990	8.845877	-21.119981
1.0980	8.869286	-25.700472
1.0970	8.897287	-30.305032
1.0960	8.929904	-34.933738
1.0950	8.967163	-39.586670
1.0940	9.009086	-44.263907
1.0930	9.055699	-48.965528
1.0920	9.107026	-53.691612
1.0910	9.163091	-58.442238
1.0900	9.223919	-63.217486

Period of Vibration.

The period of vibration (time taken for the oscillator to undertake one complete vibration, returning to its original position and velocity) can be calculated in three ways. Most direct is the calculation from the energy curve; using the gradient constitutes a faster, albeit less direct, method, while calculating it from the vibrational frequency is very fast but assumes that the vibrational spectrum has already been calculated.

(1) From the energy curve.

For a simple harmonic oscillator the period 'r' is given by

$$r = 2*pi*sqrt(m/k)$$

where m = reduced mass and k = force-constant. The reduced mass (in AMU) of a nitrogen molecule is 14.0067/2 = 7.00335, and the force-constant can be calculated from

$$E - c = 1/2 * k(R-Ro)^2 .$$

Given Ro = 1.1038, R = 1.092, c = 8.79666 and E = 9.1070Kcal/mol then

$$k = 4453 \text{ Kcal/mole/A}^2$$

$$= 4453 * 4.184 * 10^3 * 10^7 * 10^8 * 10^8 \text{ ergs/cm}^2$$

$$= 1.8631 * 10^{30} \text{ ergs/cm}^2$$

Therefore, $r = 2 * 3.14159 * sqrt(7.0035/(1.8631*10^{30})) $ seconds

$$= 12.182 * 10^{-15} \text{ seconds}$$

$$= 12.182 \text{ fS (Femtoseconds)}$$

(2) From the gradient curve.

The force constant is the derivative of the gradient W.R.T. distance

$$k = dG/dx$$

Since we are using discrete points, the force constant is best obtained from finite differences:

$$k = (G2-G1)/(x2-x1)$$

For x2 = 1.1100, G2 = 27.699 and for x1 = 1.0980, G1 = -25.700, giving rise to k = 4450 kcal/mole/A/A and a period of 12.186 fS.

(3) From the vibrational frequency.

Given a "frequency" of vibration of N2 of v=2740.36 cm(-1) the period of oscillation is given directly by

$$r = 1/(v*c)$$

$$= 1/(2740.36 * 2.998 * 10^{10}) \text{ seconds}$$

$$= 12.172 \text{ fS}$$

- 88 -

Summarizing, by three different methods the period of oscillation of N2 is calculated to be 12.182, 12.186 and 12.172 fS, average 12.180fS.

. Initial Dynamics of N2 Molecule with N-N distance = 1.1085945 Angstroms

A useful check on the dynamics of N2 is to calculate the initial acceleration of the two nitrogen atoms after releasing them from a starting interatomic separation of 1.1085945 Angstroms. This distance was chosen in order to have a starting H.o.F. of 8.8500Kcal/mole.

At $R(N-N) = 1.1085945$ Angstroms,

$$G = -21.619 \text{ Kcal/mole/Angstrom}$$
$$= -9.0454 * 10^{19} \text{ ergs/cm}$$

Therefore acceleration $f = -9.0454 * 10^{19}/14.0067 \text{ cm/sec/sec}$
$$= -6.4579 * 10^{18} \text{ cm/sec/sec}$$
$$= -6.589 * 10^{15} * \text{ Earth surface gravity!}$$

Distance from equilibrium = 0.00491 Angstroms.

After 0.1 fS velocity $= 0.1 * 10^{-15} * (-6.4579 * 10^{18}) \text{ cm/sec}$

$$= 645.79 \text{ cm/sec}$$

In the DRC the time-interval between points calculated is a complicated function of the curvature of the local surface. For Version 3.11, the first time-interval is 0.105fS, so the calculated velocity at this time should be 0.105 * 6457.9 = 678.1cm/sec, in the DRC calculation the predicted velocity is 678.1cm/sec.

. The option is provided to allow sampling of the system at constant time-intervals, the default being 0.1fS. For the first few points the calculated velocities are as follows.

TIME	CALCULATED VELOCITY	TIME	CALCULATED VELOCITY	LINEAR VELOCITY	DIFF.
0.000	0.0	0.000	0.0	0.0	0.0
0.100	645.7			645.8	-0.1
		0.105	678.1	678.1	0.0
0.200	1290.5			1291.6	-1.1
		0.215	1388.5	1388.4	+0.1
0.300	1930.5			1937.4	-6.9
		0.331	2138.6	2137.6	+1.0
0.400	2565.6			2583.2	-17.6
		0.453	2897.9	2925.4	-27.5
0.500	3193.7			3228.9	-35.2
		0.580	3693.9	3745.6	-51.7
0.600	3816.2			3874.7	-58.5

Between 0.000 and 0.105 there is a negative difference between the predicted velocity assuming v=ft and the interpolated velocity calculated by fitting the calculated velocities for the first three points using a quadratic fit. This difference arises from the decreasing acceleration reducing the value of the calculated velocity at point 3 (predicted: 1388.6, calculated: 1388.4). In order to get a smooth fit to these points, the interpolated velocity at time 0.100fS rises 0.1cm/sec above the predicted curve. The effect of deceleration on a system appears only at the second point calculated, as a result the first point calculated incurs an error (although in this example the error is negligable, less than 0.00001 Kcal/mole) due to neglect of deceleration forces. As the calculated velocity is a fourth-order polynomial of the acceleration, and the acceleration, its first, second and third derivatives, are all changing, the predicted velocity rapidly becomes a poor guide to future velocities.

For simple harmonic motion the velocity at any time is given by

$$v = v0 * sin(2*pi*t/r)$$

For a kinetic energy of 0.05335kcal/mole, each nitrogen atom has a velocity of

$$v0 = sqrt(0.05335*4.184*10^{10}/14.0067) \text{ cm/sec}$$

$$= 12623.9$$

therefore $v = 12623.9 * sin(0.51586*t)$

This gives a better fit to the calculated points.

Time	Calculated Velocity	Simple Harmonic Velocity	Difference	Best Harmonic Fit (t=12.275)	Diff.
0.000	0.0	0.0	0.0	0.0	0.0
0.100	645.7	650.9	-5.2	645.9	-0.2
0.200	1290.5	1300.1	-9.6	1290.0	+0.5
0.300	1930.5	1945.8	-15.3	1930.8	-0.3
0.400	2565.6	2586.4	-20.8	2566.6	-1.0
0.500	3193.7	3220.0	-26.3	3195.6	-1.9
0.600	3816.2	3845.1	-28.9	3816.2	0.0

The repeat-time required for this motion is 12.171 fS, in good agreement with the three values calculated using static models. The repeat time should not be calculated from the time required to go from a minimum to a maximum and then back to a minimum -- only half a cycle. For all real systems the potential energy is a skewed parabola, so that the potential energy slopes are different for both sides. Only the addition of the two half-cycles is meaningful.

Conservation of Normal Coordinate

So far this analysis has only considered a homonuclear diatomic. A detailed analysis of a large polyatomic is impractical, and for simplicity a molecule of formaldehyde will be studied.

In polyatomics, energy can transfer between modes. This is a result of the non-parabolic nature of the potential surface. For small displacements the surface can be considered as parabolic. This means that for small displacements interconversion between modes should occur only very slowly. Of the six normal modes, mode 1, at 1204.5 cm(-1), the in-plane C-H asymmetric bend is the most unsymmetric vibration, and is chosen to demonstrate conservation of vibrational purity.

Mode 1 has a frequency corresponding to 3.44 Kcal/mole and a predicted vibrational time of 27.69fS. By direct calculation, using the DRC, the cycle time is 27.55fS. The rate of decay of this mode is very fast, having an estimated half-life of only a few thousands femtoseconds.

Rate of Decay of Starting Mode

For trajectories initiated by an IRC=n calculation, whenever the potential energy is a minimum the current velocity is compared with the supplied velocity. The square of the cosine of the angle between the two velocity vectors is a measure of the intensity of the original mode in the current vibration.

Half-Life for Decay of Initial Mode

Vibrational purity is assumed to decay according to zero'th order kinetics. The half-life is thus $-0.6931472*t/log(psi2)$ fS, where psi2 is the square of the overlap integral of the original vibration with the current vibration. Due to the very slow rate of decay of the starting mode, several half-life calculations should be examined. Only when successive half-lives are similar should any confidence be placed in their value.

DRC Print Options

The amount of output in the DRC is controlled by three sets of options. These sets are:

(a) Equivalent Keywords H-PRIORITY, T-PRIORITY, and X-PRIORITY

(b) Potential Energy Turning Point option.

(c) Geometry Maxima Turning Point options.

If T-PRIORITY is used then turning points cannot be monitored. Currently H-PRIORITY and X-PRIORITY are not implemented, but will be as soon as practical.

To monitor geometry turning points, put a "T" in place of the geometry optimization flag for the relevant geometric variable.

To monitor the potential energy turning points, put a "T" for the flag for atom 1 bond length (Do not forget to put in a bond-length (zero will do)!).

The effect of these flags together is as follows.

1.   No options: All calculated points will be printed. No turning points will be calculated.

2.   Atom 1 bond length flagged with a "T": If T-PRIO, etc. are NOT specified, then potential energy turning points will be printed.

3. Internal coordinate flags set to "T": If T-PRIO, etc. are NOT
    specified, then geometry extrema will be printed. If only one
    coordinate is flagged, then the turning point will be displayed
    in chronologic order; if several are flagged then all turning
    points occuring in a given time-interval will be printed as
    they are detected. In other words, some may be out of
    chronologic order. Note that each coordinate flagged will give
    rise to a different geometry: minimize flagged coordinates to
    minimize output.

4. Potential and geometric flags set: The effect is equivalent to
    the sum of the first two options.

5. T-PRIO set: No turning points will be printed, but constant
    time-slices (by default 0.1fS) will be used to control the
    print.

6.11 SPARKLES

Four extra "elements" have been put into MOPAC. These represent
pure ionic charges, roughly equivalent to the following chemical
entities:

Chemical Symbol	Equivalent to
+	Tetramethyl ammonium radical, Potassium atom or Cesium atom.
++	Barium atom.
−	Borohydride radical, Halogen, or Nitrate radical
−−	Sulfate, oxalate.

For the purposes of discussion these entities are called
"sparkles": the name arises from consideration of their behavior.

Behavior of sparkles in MOPAC.

Sparkles have the following properties:

1. Their nuclear charge is integer, and is +1, +2, -1, or -2;
    there are an equivalent number of electrons to maintain
    electroneutrality, 1, 2, -1, and -2 respectively. For example,
    a '+' sparkle consists of a unipositive nucleus and an
    electron. The electron is donated to the quantum mechanics
    calculation.

2. They all have an ionic radius of 0.7 Angstroms. Any two
    sparkles of opposite sign will form an ion-pair with a
    interatomic separation of 1.4A.

3. They have a zero heat of atomization, no orbitals, and no ionization potential.

They can be regarded as unpolarizable ions of diameter 1.4A. They do not contribute to the orbital count, and cannot accept or donate electrons.

Since they appear as uncharged species which immediately ionize, attention should be given to the charge on the whole system. For example, if the alkaline metal salt of formic acid was run, the formula would be:

HCOO+ where + is the unipositive sparkle. The charge on the system would then be zero.

A water molecule polarized by a positive sparkle would have the formula H2O+, and the charge on the system would be +1

At first sight, a sparkle would appear to be too ionic to be a point charge and would combine with the first charge of opposite sign it encountered.

This representation is faulty, and a better description would be of an ion, of diameter 1.4A, and the charge delocalized over its surface. Computationally, a sparkle is an integer charge at the center of a repulsion sphere of form exp(-alpha*r). The hardness of the sphere is such that other atoms or sparkles can approach within about 2 Angstroms quite easily, but only with great difficulty come closer than 1.4A.

Uses of Sparkles

1. They can be used as counterions, e.g. for acid anions or for cations. Thus, if the ionic form of an acid is wanted, then the moieties H.X, H.-, and +.X could be examined.

2. Two sparkles of equal and opposite sign can form a dipole for mimicking solvation effects. Thus water could be surrounded by six dipoles to simulate the solvent cage. A dipole of value D can be made by using the two sparkles + and -, or using ++ and --. If + and - are used, the inter-sparkle separation would be D/4.803 Angstroms. If ++ and -- are used, the separation would be D/9.606 Angstroms. If the inter-sparkle separation is less than 1.0 Angstroms (a situation that cannot occur naturally) then the energy due to the dipole on its own is subtracted from the total energy.

3. They can operate as polarization functions. A controlled, shaped electric field can easily be made from two or more sparkles. The polarizability in cubic Angstroms of a molecule in any particular orientation can then easily be calculated.

6.12 MECHANISM OF THE FRAME IN THE FORCE CALCULATION

The FORCE calculation uses cartesian coordinates, and all 3N modes are calculated, where N is the number of atoms in the system. Clearly, there will be 5 or 6 "trivial" vibrations, which represent the three translations and two or three rotations. If the molecule is exactly at a stationary point, then these "vibrations" will have a force constant and frequency of precisely zero. If the force calculation was done correctly, and the molecule was not exactly at a stationary point, then the three translations should be exactly zero, but the rotations would be non-zero. The extent to which the rotations are non-zero is a measure of the error in the geometry.

If the distortions are non-zero, the trivial vibrations can interact with the low-lying genuine vibrations or rotations, and with the transition vibration if present.

To prevent this the analytic form of the rotations and vibrations is calculated, and arbitrary eigenvalues assigned; these are 500, 600, 700, 800, 900, and 1000 millidynes/angstrom for Tx, Ty, Tz, Rx, Ry and Rz (if present), respectively. The rotations are about the principal axes of inertia for the system, taking into account isotopic masses. The "force matrix" for these trivial vibrations is determined, and added on to the calculated force matrix. After diagonalization the arbitrary eigenvalues are subtracted off the trivial vibrations, and the resulting numbers are the "true" values. Interference with genuine vibrations is thus avoided.

6.13 PSEUDODIAGONALIZATION -SUBROUTINE DIAG

The basis of subroutine DIAG is the observation that accurate matrix diagonalization of the secular determinant is not a prerequisite in the SCF procedure for obtaining a self-consistent density matrix in a variationally optimized calculation. To have a self-consistent density matrix it is sufficient to have annihilated all energy matrix elements connecting the occupied and virtual molecular orbitals.

<div align="center">THEORY</div>

Given a basis set of N atomic orbitals and Ne electrons, there will be No = Ne /2 occupied molecular orbitals and Nvir = N-Nv virtual orbitals. If the approximate form of the molecular orbitals is known, perhaps from an accurate, standard diagonalization of the first trial secular determinant, then the interaction matrix can be constructed. The off-diagonal matrix elements in Fov can then be annihilated by a series of 2x2 rotations in the manner of Jacobi. Unlike Jacobi, however, the method need not be cycled to exactly diagonalize Fov; it is sufficient to have only one sweep. This is due to the fact that the Fock equations form a pseudo-eigenvalue problem, and it is necessary to iterate to obtain a self-consistent field. For the same reason the second-order effects of the 2x2 rotations can be ignored. In an exact diagonalization the off-diagonal matrix elements formed by an elementary 2 by 2 rotation would have to be eliminated. These are normally less

than one tenth of the matrix element being annihilated, and as the SCF procedure does not converge at one magnitude per cycle the second-order errors introduced can be absorbed into the Fock matrix of the following cycle. Also, since second-order effects in the "diagonalization" are being ignored it is equally valid to eliminate only those matrix elements which are comparable with the largest off-diagonal elements in Fov. A further advantage of the pseudo-eigenvalue nature of the SCF equations appears when we come to evaluate the diagonal terms of the secular determinant. For this, we can equate these elements with the eigenvalues resulting from the exact diagonalization, and hold them exactly constant throughout the entire calculation, right up to self-consistency. At first sight this would appear to introduce errors in the final SCF density matrix, as obviously the sum of the eigenvalues cannot be constant in an exact calculation, and thus the final sum of occupied energy levels must be in error. However, to obtain a SCF density matrix we not only do not need to know the exact eigenvalues, we have no need to know the sum of the occupied energy levels. Using the initial set of eigenvalues, a 2 by 2 rotation will, of course, not eliminate fully even those elements which we do choose to operate on, but again the pseudo-eigenvalue nature of the problem comes to our rescue. As the iterations proceed those errors introduced are rapidly eliminated, so that at self-consistency an exact density matrix is generated. But we have no knowledge of the values of the eigenvalues, eigenvectors or two-electron energy. This completes the definition of the secular determinant, and the 2 by 2 rotations needed to pseudo-diagonalize it. The unitary matrix that results from the set of rotations would normally need to be multiplied by the original set of eigenvectors to obtain the correct molecular orbital matrix. There is no reason to start the pseudo-diagonalization with a unit matrix, and this step can be eliminated by starting the pseudo-diagonalization using the old set of molecular orbitals.

COMPUTATIONAL ASPECTS

Clearly, an initial exact diagonalization is needed in order to obtain a good starting set of eigenvectors. This is, however, not a sufficient condition for initiating the use of the new method. During the first one or two cycles of a SCF procedure the order of occupancy of the molecular orbitals may change. That is, the occupancy of the M.O.'s in the first two iterations may correspond to an excited singlet state. For this reason it is recommended that the new method be used only after the initial large fluctuations in the density matrix have died down. Our arbitrary criterion was that the largest change in the diagonal elements of the density matrix between any two iterations should be less than 0.05.

The method proposed uses three matrices: the C matrix of eigenvector coefficients, the F matrix, and a working matrix to hold the new Fov matrix. In virtual memory computers this would normally not present any problem. A criterion is also needed to decide when to perform a 2 by 2 rotation and elimination. After a few trial calculations, it was found to be efficient to eliminate all off-diagonal elements whose modulus was larger than 0.01 times the modulus of the largest off-diagonal element. No important change in the number of iterations was observed. The electronic energy calculated after each iteration does indeed differ following the introduction of the new

method. This is a result of all the second-order effects being
introduced, b.. the differences between electronic energies calculated
by the exact diagonalization and by the new method rapidly converge.   An
interesting side-effect of the new non-rigorous method is that some
damping is introduced: in all molecules examined the number of
iterations required to achieve self-consistency either stayed constant
or dropped by one or two.

6.14 DYNAMIC REACTION COORDINATE

Introduction

The course of a molecular vibration can be followed by calculating the potential and kinetic energy at various times. Two extreme conditions can be identified: (a) gas phase, in which the total energy is a constant through time, there being no damping of the kinetic energy allowed, and (b) liquid phase, in which kinetic energy is always set to zero, the motion of the atoms being infinitely damped.

All possible degrees of damping are allowed. In addition, the facility exists to dump energy into the system, appearing as kinetic energy. As kinetic energy is a function of velocity, a vector quantity, the energy appears as energy of motion in the direction in which the molecule would naturally move. If the system is a transition state, then the excess kinetic energy is added after the intrinsic kinetic energy has built up to at least 0.2Kcal/mole.

For ground-state systems, the excess energy sometimes may not be added; if the intrinsic kinetic energy never rises above 0.2kcal/mole then the excess energy will not be added.

Equations Used

Force acting on any atom

g(i) + g'(i)t + g"(i)t**2 = dE/dx(i) + d**2E/dx(i)**2 +d**3E/dx(i)**3

Acceleration due to force acting on each atom

a(i) = (g(i)+g'(i)t+g"(i)t**2)/M(i)

New velocity

V(o)+Dt*g(i)/M(i)+1/2*Dt**2*g'(i)/M(i)+/3*Dt**3*g"(i)/M(i)

or

V(i) = V(i) + V'(i)t + V''(i)t**2 + V'''(i)t**3

That is, the change in velocity is equal to the integral over the time interval of the acceleration.

New position of atoms

X(i) = X(o) + V(o)t + 1/2*V't**2 + 1/3*V''t**3 + 1/4*V'''t**4

That is, the change in position is equal to the integral over the time interval of the velocity.

The velocity vector is accurate to the extent that it takes into account the previous velocity, the current acceleration, the predicted acceleration, and the change in predicted acceleration over the time interval. Very little error is introduced due to higher order contributions to the velocity; those that do occur are absorbed in a re-normalization of the magnitude of the velocity vector after each time interval.

The magnitude of Dt, the time interval, is determined mainly by the factor needed to re-normalize the velocity vector. If it is significantly different from unity, Dt will be reduced; if it is very close to unity, Dt will be increased.

Even with all this, errors creep in and a system, started at the transition state, is unlikely to return precisely to the transition state unless an excess kinetic energy is supplied, for example 0.2Kcal/mole.

The calculation is carried out in cartesian coordinates, and converted into internal coordinates for display. All cartesian coordinates must be allowed to vary, in order to conserve angular and translational momentum.

6.15 CONFIGURATION INTERACTION

MOPAC contains a very large Multi-Electron Configuration Interaction calculation, MECI, which allows almost any configuration interaction calculation to be performed. Because of its complexity, two distinct levels of input are supported; the default values will be of use to the novice while an expert has available an exhaustive set of keywords from which a specific C.I. can be tailored.

A MECI calculation involves the interaction of microstates representing specific permutations of electrons in a set of M.O.'s. Starting with a set electronic configuration, either closed shell or open shell, but unconditionally restricted Hartree-Fock, the first step in a MECI calculation is the removal from the M.O.'s of the electrons to be used in the C.I.

Each microstate is then constructed from these empty M.O.'s by adding in electrons according to a prescription. The energy of the configuration is evaluated, as is the energy of interaction with all previously-defined configurations. Diagonalization then results in state functions. From the eigenvectors the expectation value of s**2 is calculated, and the spin-states of the state functions calculated.

General Overview of Keywords

Keywords associated with the operations of MECI are:

SINGLET	DOUBLET	EXCITED
TRIPLET	QUARTET	BIRADICAL
QUINTET	SEXTET	ESR
OPEN(nl,n2)	C.I.=n	MECI
ROOT=n		

Each keyword may imply others; thus TRIPLET implies an open-shell system, therefore OPEN(2,2), and C.I.=2 are implied, if not user specified.

Starting Electronic Configuration

MECI is restricted to RHF calculations, but with that single restriction any starting configuration will be supported. Examples of starting configurations would be

System	KeyWords used	Starting Configuration
Methane	<none>	2.00 2.00 2.00 2.00 2.00
Methyl Radical	<none>	2.00 2.00 2.00 2.00 1.00
Twisted Ethylene	TRIPLET	2.00 2.00 2.00 1.00 1.00
Twisted Ethylene	OPEN(2,2)	2.00 2.00 2.00 1.00 1.00
Twisted Ethylene Cation	OPEN(1,2)	2.00 2.00 2.00 0.50 0.50
Methane Cation	CHARGE=1 OPEN(5,3)	2.00 2.00 1.67 1.67 1.67

Choice of starting configuration is important. For example, if twisted ethylene, a ground-state triplet, is not defined using TRIPLET or OPEN(2,2), then the closed-shell ground-state structure will be calculated. Obviously, this configuration is a legitimate microstate, but from the symmetry of the system a better choice would be to define one electron in each of the two formally degenerate pi-type M.O.'s. The initial SCF calculation does not distinguish between OPEN(2,2) and TRIPLET since both keywords define the same starting configuration. This can be verified by monitoring the convergence using PL, for which both keywords give the same SCF energy.

Removal of Electrons from Starting Configuration

For a starting configuration of alpha M.O. occupancies $O(i)$, $O(i)$ being in the range 0.0 to 1.0, the energies of the M.O.'s involved in the MECI can be calculated from

$$E(i) = Sum(j)(2J(i,j)-K(i,j))O(j)$$

where $J(i,j)$ and $K(i,j)$ are the coulomb and exchange integrals between M.O.'s i and j. The M.O. index j runs over those M.O.'s involved in the MECI only. Most MECI calculations will involve between 1 and 5

M.O.'s, so a system with about 30 filled or partly filled M.O.'s could have M.O.'s 25-30 involved. The resulting eigenvalues correspond to those of the cationic system resulting from removal of n electrons, where n is twice the sum of the orbital occupancies of those M.O.'s involved in the C.I.

The arbitrary zero of energy in a MECI calculation is the starting ground state, without any correction for errors introduced by the use of fractional occupancies. In order to calculate the energy of the various configurations, the energy of the vacuum state (i.e., the state resulting from removal of the electrons used in the C.I.) needs to be evaluated. This energy is defined by

$$GSE = Sum(i)[\ E(i)O(i) + J(i,i) * O(i)*O(i)$$
$$+ Sum(j<i)(\ 2(2J(i,j) - K(i,j)\ ) * O(i)*O(j)\ )\ ]$$

Formation of Microstate Configuration

Microstates are particular electron configurations. Thus if there are 5 electrons in 5 levels, then various microstates could be as follows:

Microstates for 5 electrons in 5 M.O.'s

	Electron Configuration				Electron Configuration		
	Alpha	Beta	M(s)		Alpha	Beta	M(s)
	1 2 3 4 5	1 2 3 4 5			1 2 3 4 5	1 2 3 4 5	
1	1,1,1,0,0	1,1,0,0,0	1/2	4	1,1,1,1,1	0,0,0,0,0	5/2
2	1,1,0,0,0	1,1,1,0,0	-1/2	5	1,1,0,1,0	1,1,0,0,0	1/2
3	1,1,1,0,0	0,0,0,1,1	1/2	6	1,1,0,1,0	1,0,1,0,0	1/2

For 5 electrons in 5 M.O.'s there are 252 microstates (10!/(5!*5!)), but as states of different spin do not mix, we can use a smaller number. If doublet states are needed then 100 states (5!/(2!*3!)*(5!/3!*2!) are needed. If only quartet states are of interest then 25 states (5!/(1!*4!)*(5!/4!*1!) are needed and if the sextet state is required, then only one state is calculated.

In the microstates listed, state 1 is the ground-state configuration. This can be written as (2,2,1,0,0), meaning that M.O.'s 1 and 2 are doubly occupied, M.O. 3 is singly occupied by an alpha electron, and M.O.'s 4 and 5 are empty. Microstate 1 has a component of spin of 1/2, and is a pure doublet. By Kramer's degeneracy – sometimes called time-inversion symmetry – microstate 2 is also a doublet, and has a spin of 1/2 and a component of spin of -1/2.

Microstate 3, while it has a component of spin of 1/2, is not a doublet, but is in fact a component of a doublet, a quartet and a sextet. The coefficients of these states can be calculated from the Clebsch-Gordon 3-J symbol. For example, the coefficient in the sextet is 1/Sqrt(5).

Microstate 4 is a pure sextet. If all 100 microstates of component of spin = 1/2 were used in a C.I., one of the resulting states would have the same energy as the state resulting from microstate 4.

Microstate 5 is an excited doublet, and microstate 6 is an excited state of the system, but not a pure spin-state.

By default, if n M.O.'s are included in the MECI, then all possible microstates which give rise to a component of spin = 0 for even electron systems, or 1/2 for odd electron systems, will be used.

Permutations of Electrons among Molecular Orbitals

(0,1) =	0	(2,4) =	1100	(3,5) =	11100	(2,5) =	11000
			1010		11010		10100
(1,1) =	1		1001		11001		10010
			0110		10110		10001
(0,2) =	0		0101		10101		01100
			0011		10011		01010
(1,2) =	10				01110		01001
	01	(1,4) =	1000		01101		00110
			0100		01011		00101
(1,3) =	100		0010		00111		00011
	010		0001				
	001						
(2,3) =	110						
	101						
	011						

Sets of Microstates for Various MECI Calculations

	Odd Electron Systems			Even Electron Systems		
	Alpha	Beta	No. of Configs.	Alpha	Beta	No. of Configs.
C.I.=1	(1,1) *	(0,1)	= 1	(1,1) *	(1,1)	= 1
2	(1,2) *	(0,2)	= 2	(1,2) *	(1,2)	= 4
3	(2,3) *	(1,3)	= 9	(2,3) *	(2,3)	= 9
4	(2,4) *	(1,4)	= 24	(2,4) *	(2,4)	= 36
5	(3,5) *	(2,5)	= 100	(3,5) *	(3,5)	= 100

Multi Electron Configurat'on Interaction

The numbering of the M.O.'s used in the MECI is standard, and follows the Aufbau principle. The order of filling is in order of energy, and alpha before beta. This point is critically important in deciding the sign of matrix elements. For a 5 M.O. system, then, the order of filling is.

$$(1)(\bar{1})(2)(\bar{2})(3)(\bar{3})(4)(\bar{4})(5)(\bar{5})$$

A triplet state arising from two microstates, each with a component of spin = 0, will thus be the positive combination.

$$(\bar{1})(2) \quad + \quad (1)(\bar{2})$$

This is in variance with the sign convention used in earlier programs for running MNDO. This standard sign convention was chosen in order to allow the signs of the microstate coefficients to conform to those resulting from the spin step-down operator.

Matrix elements between all pairs of microstates are calculated in order to form the secular determinant. Many elements will be identically zero, due to the interacting determinants differing by more than two M.O.'s. For the remaining interactions the following types can be identified.

1. The two determinants are identical:

    No permutations are necessary in order to calculate the sign of the matrix element. E(p,p) is given simply by

    $$E(p,p)=Sum(i)Oa(i,p)*[Eig(i) + Sum(1/2(<ii|jj>-<ij|ij>)*Oa(j,p)$$
    $$+ \qquad <ii|jj> \qquad *Ob(j,p))$$

    $$+ Sum(i)Ob(i,p)*[Eig(i) + Sum(1/2(<ii|jj>-<ij|ij>)*Ob(j,p))$$

    Oa(i,p) = Occupancy of alpha M.O. i in Microstate p
    Ob(i,p) = Occupancy of beta M.O. i in Microstate p

2. Determinants differing by exactly one M.O.:

    The differing M.O. can be of type alpha or beta. It is sufficient to evaluate the case in which both M.O.'s are of alpha type, the beta form is obtained in like manner.

    $$E(p,q) = Sum(k) [<ij|kk> - <ik|jk>) * (Occa(k) - Occg(k))$$
    $$+ <ij|kk> \qquad * (Occb(k) - Occg(k)]$$

E(p,q) may need to be multiplied by -1, if the number of
two electron permutations required to bring M.O.'s i and j
into coincidence is odd.

Where Occa(k) is the alpha molecular orbital occupancy in
the configuration interaction.

3. Determinants differing by exactly two M.O.'s:

The two M.O.'s can have the same or opposite spins. Three
cases can be identified:

1. Both M.O.'s have alpha spin:

For the first microstate having M.O.'s i and j, and
the second microstate having M.O.'s k and l, the matrix
element connecting the two microstates is given by

Q(p,q) = <ik|jl> - <il|jk>
E(p,q) may need to be multiplied by -1, if the number of
two electron permutations required to bring M.O. i into
coincidence with M.O. k and M.O. j into coincidence with
M.O. l is odd.

2. Both M.O.'s have beta spin:

The matrix element is calculated in the same manner as
in the previous case.

3. One M.O. has alpha spin, and one beta spin:

For the first microstate having M.O.'s alpha(i) and
beta(j), and the second microstate having M.O.'s alpha(k)
and beta(l), the matrix element connecting the two
microstates is given by

Q(p,q) = <ik|jl>

E(p,q) may need to be multiplied by -1, if the number of
two electron permutations required to bring M.O. i into
coincidence with M.O. k and M.O. j into coincidence with
M.O. l is odd.

States Arising from Various Calculations

Each MECI calculation invoked by use of the keyword C.I.=n normally gives rise to states of quantized spins. When C.I. is used without any other modifying keywords, the following states will be obtained.

No. of M.O.'s in MECI	States Arising From Odd Electron Systems Doublets			States Arising From Even Electron Systems Singlets	Triplets	
1	1			1		
2	2			3	1	
3	8	1		6	3	
4	20	4		20	15	1
5	75	24	1	50	45	5

These numbers of spin states will be obtained irrespective of the chemical nature of the system.

Calculation of Spin-States

In order to calculate the spin-state, the expectation value of S2 is calculated.

$$S2 = S(S+1) = Sz**2 + 2*S(+)S(-)$$

$$= Ne -$$
$$Sum(i) [C(i,k)*C(i,k)*(1/4*(Na(i)-Nb(i))**2$$
$$+ Sum(1) Oa(1,i)*Ob(1,i))$$
$$+Sum(j) 2[C(i,k)*C(j,k)*(Kronekerdelta(C(i,k)(S(+)S(-))C(j,k)]]$$

Where Ne   = No. of electrons in C.I.
       C(i,k) = Coefficient of Microstate i in State k
       Na(i)   = Number of alpha electrons in Microstate i
       Nb(i)   = Number of beta electrons in Microstate i
       Oa(1,k) = Occupancy of alpha M.O. 1 in Microstate k
       Ob(1,k) = Occupancy of beta M.O. 1 in Microstate k
       S(+)   = Spin shift up or step up operator
       S(-)   = Spin shift down or step down operator
       The Kronekerdelta is 1 if the two terms in brackets following it are identical.

The spin state is calculated from $S = 1/2 (Sqrt(1+4*S2) - 1)$

In practice, S is calculated to be exactly integer, or half integer. That is, there is insignificant error due to approximations used. This does not mean, however, that the method is accurate. The spin calculation is completely precise, in the group theoretic sense,

but the accuracy of the calculation is limited by the Hamiltonian used,
a space-dependent function.

Choice of State to be Optimized

MECI can calculate a large number of states of various total spin.
Two schemes are provided to allow a given state to be selected. First,
ROOT=n will, when used on its own, select the n'th state, irrespective
of its total spin. By default n=1. If ROOT=n is used in conjunction
with a keyword from the set SINGLET, DOUBLET, TRIPLET, QUARTET, QUINTET,
or SEXTET, then the n'th root of that spin-state will be used. For
example, ROOT=4 and SINGLET will select the 4th singlet state. If there
are two triplet states below the fourth singlet state then this will
mean that the sixth state will be selected.

Calculation of Unpaired Spin Density

Starting with the state functions as linear combinations of
configurations, the unpaired spin density, corresponding to the alpha
spin density minus the beta spin density, will be calculated for the
first few states. This calculation is straightforward for diagonal
terms, and only those terms are used.

6.16 REDUCED MASSES IN A FORCE CALCULATION

Reduced masses for a diatomic are given by

$$\frac{(mass1) * (mass2)}{(mass1) + (mass2)}$$

For a Hydrogen molecule the reduced mass is thus 0.5; for heavily hydrogenated systems, e.g. methane, the reduced mass can be very low. A vibration involving only heavy atoms , e.g. a C-N in cyanide, should give a large reduced mass.

For the "trivial" vibrations the reduced mass is ill-defined, and where this happens the reduced mass is set to zero.

6.17 USE OF SADDLE CALCULATION

A SADDLE calculation uses two complete geometries, as shown on the following data file for the ethyl radical hydrogen migration from one methyl group to the other.

```
Line  1:              UHF   SADDLE
Line  2:         ETHYL RADICAL HYDROGEN MIGRATION
Line  3:
Line  4:    C    0.000000 0      0.000000 0      0.000000 0    0  0  0
Line  5:    C    1.479146 1      0.000000 0      0.000000 0    1  0  0
Line  6:    H    1.109475 1    111.328433 1      0.000000 0    2  1  0
Line  7:    H    1.109470 1    111.753160 1    120.288410 1    2  1  3
Line  8:    H    1.109843 1    110.103163 1    240.205278 1    2  1  3
Line  9:    H    1.082055 1    121.214083 1     38.110989 1    1  2  3
Line 10:    H    1.081797 1    121.521232 1    217.450268 1    1  2  3
Line 11:    O    0.000000 0      0.000000 0      0.000000 0    0  0  0
Line 12:    C    0.000000 0      0.000000 0      0.000000 0    0  0  0
Line 13:    C    1.479146 1      0.000000 0      0.000000 0    1  0  0
Line 14:    H    1.109475 1    111.328433 1      0.000000 0    2  1  0
Line 15:    H    1.109470 1    111.753160 1    120.288410 1    2  1  3
Line 16:    H    2.109843 1     30.103163 1    240.205278 1    2  1  3
Line 17:    H    1.082055 1    121.214083 1     38.110989 1    1  2  3
Line 18:    H    1.081797 1    121.521232 1    217.450268 1    1  2  3
Line 19:    O    0.000000 0      0.000000 0      0.000000 0    0  0  0
Line 20:
```

Details of the mathematics of SADDLE appeared in print in 1984, (M. J. S. Dewar, E. F. Healy, J. J. P. Stewart, J. Chem. Soc. Faraday Trans. II, 3, 227, (1984)) so only a superficial description will be given here.

The main steps in the saddle calculation are as follows:

1. The heats of formation of both systems are calculated.

2. A vector R of length 3N-6 defining the difference between the two geometries is calculated.

3. The scalar P of the difference vector is reduced by some fraction, normally about 5 to 15 percent.

4. Identify the geometry of lower energy; call this G.

5. Optimize G, subject to the constraint that it maintains a constant distance P from the other geometry.

6. If the newly-optimized geometry is higher in energy then the other geometry, then go to 1. If it is higher, and the last two steps involved the same geometry moving, make the other geometry G without modifying P, and go to 5.

7. Otherwise go back to 2.

The mechanism of 5 involves the coordinates of the moving geometry being perturbed by an amount equal to the product of the discrepancy between the calculated and required P and the vector R.

As the specification of the geometries is quite difficult, in that the difference vector depends on angles (which are, of necessity ill-defined by 360 degrees) SADDLE can be made to run in cartesian coordinates using the keyword XYZ. If this option is chosen then the initial steps of the calculation are as follows:

1. Both geometries are converted into cartesian coordinates.

2. Both geometries are centered about the origin of cartesian space.

3. One geometry is rotated until the difference vector is a minimum - this minimum is within 1 degree of the absolute bottom.

4. The SADDLE calculation then proceeds as described above.

LIMITATIONS:

The two geometries must be related by a continuous deformation of the coordinates. By default, internal coordinates are used in specifying geometries, and while bond lengths and bond angles are unambiguously defined (being both positive), the dihedral angles can be positive or negative. Clearly 300 degrees could equally well be specified as -60 degrees. A wrong choice of dihedral would mean that instead of the desired reaction vector being used, a completely incorrect vector was used, with disastrous results.

To correct this, ensure that one geometry can be obtained from the other by a continuous deformation, or use the XYZ option.

6.18 HOW TO ESCAPE FROM A HILLTOP

A particularly irritating phenomenon sometimes occurs when a transition state is being refined. A rough estimate of the geometry of the transition state has been obtained by either a SADDLE or reaction path or by good guesswork. This geometry is then refined by SIGMA or by NLLSQ, and the system characterized by a force calculation. It is at this point that things often go wrong. Instead of only one negative force constant, two or more are found. In the past, the recommendation has been to abandon the work and to go on to something less masochistic. It is possible, however, to systematically progress from a multiple maximum to the desired transition state. The technique used will now be described.

If a multiple maximum is identified, most likely one negative force constant corresponds to the reaction coordinate, in which case the objective is to render the other force constants positive. The associated normal mode eigenvalues are complex, but in the output are printed as negative frequencies, and for the sake of simplicity will be described as negative vibrations. Use DRAW-2 to display the negative vibrations, and identify which mode corresponds to the reaction coordinate. This is the one we need to retain.

Hitherto, simple motion in the direction of the other modes has proved difficult. However the DRC provides a convenient mechanism for automatically following a normal coordinate. Pick the largest of the negative modes to be annihilated, and run the DRC along that mode until a minimum is reached. At that point, refine the geometry once more using SIGMA and repeat the procedure until only one negative mode exists.

To be on the safe side, after each DRC+SIGMA sequence do the DRC+SIGMA operation again, but use the negative of the initial normal coordinate to start the trajectory. After both stationary points are reached, choose the lower point as the starting point for the next elimination. The lower point is chosen because the transition state wanted is the highest point on the lowest energy path connecting reactants to products. Sometimes the two points will have equal energy: this is normally a consequence of both trajectories leading to the same point or symmetry equivalent points.

After all spurious negative modes have been eliminated, the remaining normal mode corresponds to the reaction coordinate, and the transition state has been located.

This technique is relatively rapid, and relies on starting from a stationary point to begin each trajectory. If any other point is used, the trajectory will not be even roughly simple harmonic. If, by mistake, the reaction coordinate is selected, then the potential energy will drop to that of either the reactants or products, which,

incidentally, forms a handy criterion for selecting the spurious modes: if the potential energy only drops by a small amount, and the time evolution is roughly simple harmonic, then the mode is one of the spurious modes. If there is any doubt as to whether a minimum is in the vicinity of a stationary point, allow the trajectory to continue until one complete cycle is executed. At that point the geometry should be near to the initial geometry.

Superficially, a line-search might appear more attractive than the relatively expensive DRC. However, a line-search in cartesian space will normally not locate the minimum in a mode. An obvious example is the mode corresponding to a methyl rotation.

<div align="center">Keyword Sequences to be Used</div>

1. To locate the starting stationary point given an approximate transition state:-

    SIGMA

2. To define the normal modes:-

    FORCE ISOTOPE

    At this point, copy all the files to a second filename, for use later.

3. Given vibrational frequencies of -654, -123, 234, and 456, identify via DRAW-2 the normal coordinate mode, let's say that is the -654 mode. Eliminate the second mode by:

    IRC=2 DRC T=30M RESTART

    Use is made of the FORCE restart file.

4. Identify the minimum in the potential energy surface using the VAX SEARCH command, of form:

    SEARCH <Filename>.OUT %

5. Edit out of the output file the data file corresponding to the lowest point, and refine the geometry using:

    SIGMA

6. Repeat the last three steps but for the negative of the normal mode, using the copied files. The keyword for the first of the two jobs is:

    IRC=-2 DRC T=30M RESTART

7. Repeat the last four steps as often as there are spurious modes.

8. Finally, carry out a DRC to confirm that the transition state does, in fact, connect the reactants and products. The drop in potential energy should be monotonic. If you are unsure whether this last operation will work successfully, do it at any time you have a stationary point. If it fails at the very start, then we are back where we were last year -- give up and go home!!

6.19 POLARIZABILITY CALCULATION

If the electrons in a molecule are easily moved as the result of a stimulus, then the molecule is easily polarizable. Thus, if an applied electric field can easily induce a dipole, then the polarizability is large. Any induced dipole will lower the energy of the system, but this stabilization might be masked by the presence of a permanent dipole. To avoid this, use is made of an alternating electric field. If the molecule has an intrinsic dipole, then the molecule will be stabilized in one direction. When the field is reversed, the molecule will be destabilized, but, on averaging the two effects, the result is a net stabilization due only to the induced dipole.

Originally, MOPAC calculated the polarizability of molecules, radicals, and ions by use of a shaped electric field.

In the current version of MOPAC the polarizability and hyperpolarizability are calculated by direct perturbation of the Hamiltonian matrix elements. This technique was developed by Dr. Henry A. Kurtz of Memphis State University while on a USAF-UES Summer Faculty Research Program. The following discussion assumes that a homogeneous electric field gradient exists across the molecule.

The heat of formation of the molecule in this field is then calculated. This quantity can be expressed as a series sum.

Heat = H.o.F - V*E(Charge) - dV/dx*E(Dipole)

       - d2V/dx2*E(polarizability)

That is, the heat of formation in the field is the sum of the basic heat of formation, less the electric potential times any charge, any dipole times the electric field gradient, and any polarizability times the square of the electric field gradient.

We are interested in the polarizability, P.

$$P = (2/23.061)*d**2H/dE**2$$

The second derivative of H with respect to E is given by

$$d**2H/dE**2 = (H(E)+H(-E)-2*H(0))/(2*E),$$

H(E) being the heat of formation in the electric field.

The polarizability volume, Vol, is calculated from the polarizability by

$$Vol=P/(E*4*pi*E(o)) = 2/(E*23.061*4*pi*E(o)) * d**2H/dE**2$$

Substituting for E we have

$$Vol=2*1**4*pi*E(o)/(23.061*Q*Q*C*C*(1-1/2**(-1/3))**2) * d**2H/dE**2.$$

It is a simple matter to evaluate the value of this second-rank tensor by calculating the heats of formation of the molecule subject to four different electric field gradients. For the tensor component V(i,j), i=x or y or z, j=x or y or z, the directions of the four different fields are defined by.

Field 1	+i, +j	Field 2	+i, -j
Field 2	-i, -j	Field 4	-i, +j.

Thus if i=x and j=x the four fields are

Field 1	+x
Field 2	0
Field 3	-x
Field 4	0

Using these four heats of formation, in Kcal/mole, the polarizability can be calculated in units of cubic angstroms via

$$Vol = \frac{(Heat(2)+Heat(4)-Heat(1)-Heat(3))*(1*1*1*1)*2*pi*Eo}{23.061 * (1-a) * (1-a) * Q * Q * C * C}$$

1eV is 1.60219 * 10-19 Joules

Eo is 8.854188 * 10-12 Joules**(-1).C**2.M**(-1)

or 8.854188 * 10-22 Joules**(-1).C**2.Angstroms**(-1)

$$Vol = \frac{(eV * 1**4 * J**(-1) * C**2 * M**(-1))}{C**2}$$

$$Vol = 2 * 3.1415926 * 8.854188*10-22 / (23.061 * 1.60219*10-19)$$

$$=\frac{(Heat(2)+Heat(4)-Heat(1)-Heat(3))*0.0015056931*(1*1*1*1)}{(1-a)*(1-a)*Q*Q}$$

Monopolar and dipolar terms are eliminated in this treatment.

Finally, monatomic additive terms are included when MNDO is used.

A polarization matrix of size 3 * 3 is constructed and diagonalized, and the resulting eigenvalues are the calculated independent polarization volumes in cubic Angstroms; the vectors are the independent polarization vectors.

6.20 SOLID STATE CAPABILITY

Currently MOPAC can only handle up to one-dimensional extended systems. As the solid-state method used is unusual, details are given at this point.

If a polymer unit cell is large enough, then a single point in k-space, the Gamma point, is sufficient to specify the entire Brillouin zone. The secular determinant for this point can be constructed by adding together the Fock matrix for the central unit cell plus those for the adjacent unit cells. The Born-von Karman cyclic boundary conditions are satisfied, and diagonalization yields the correct density matrix for the Gamma point.

At this point in the calculation, conventionally, the density matrix for each unit cell is constructed. Instead, the Gamma-point density and one-electron density matrices are combined with a "Gamma-point-like" Coulomb and exchange integral strings to produce a new Fock matrix. The calculation can be visualized as being done entirely in reciprocal space, at the Gamma point.

Most solid-state calculations take a very long time. These calculations, called "Cluster" calculations after the original publication, require between 1.3 and 2 times the equivalent molecular calculation.

A minor "fudge" is necessary to make this method work. The contribution to the Fock matrix element arising from the exchange integral between an atomic orbital and its equivalent in the adjacent unit cells is ignored. This is necessitated by the fact that the density matrix element involved is invariably large.

The unit cell must be large enough that an atomic orbital in the center of the unit cell has an insignificant overlap with the atomic orbitals at the ends of the unit cell. In practice, a translation vector of more that about 7 or 8 Angstroms is sufficient. For one rare group of compounds a larger translation vector is needed. Polymers with delocalized pi-systems, and polymers with very small band-gaps will require a larger translation vector, in order to accurately sample k-space. For these systems, a translation vector in the order of 15-20 Angstroms is needed.
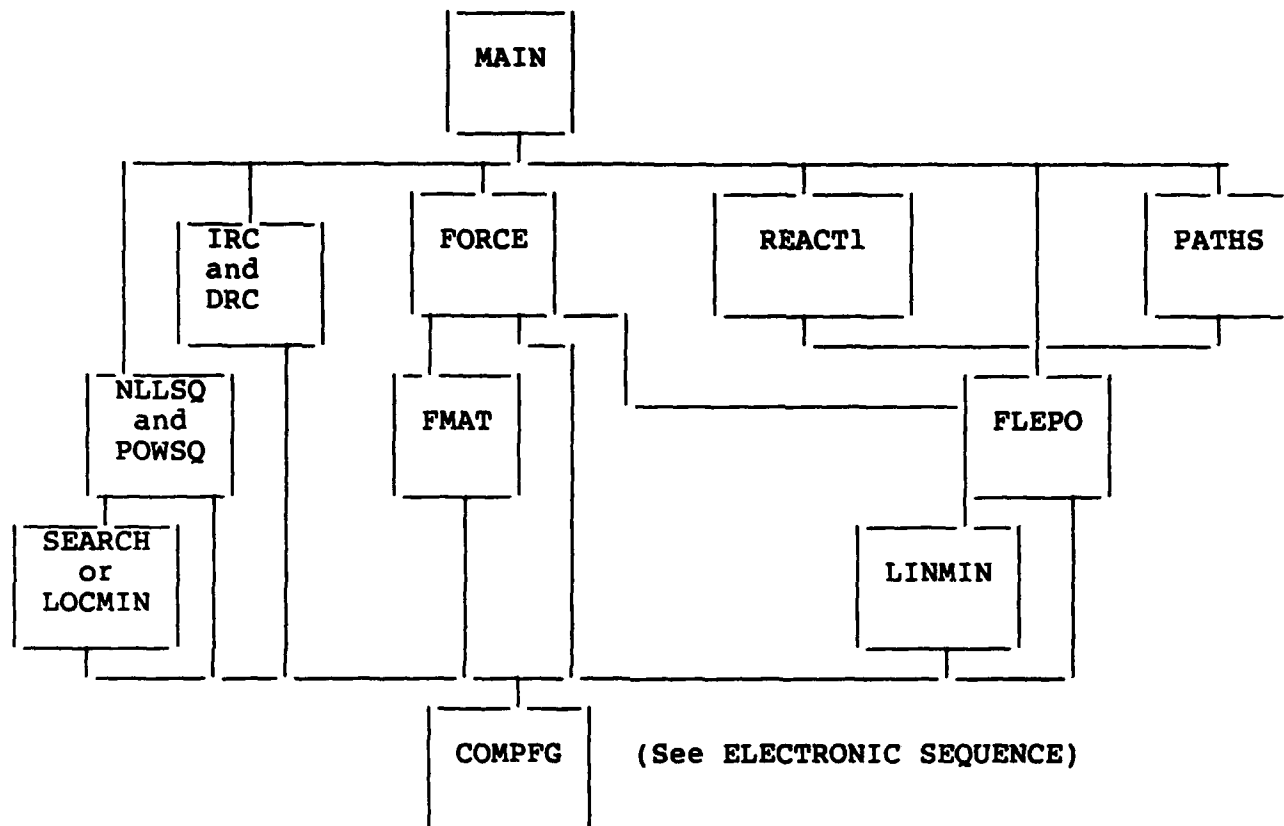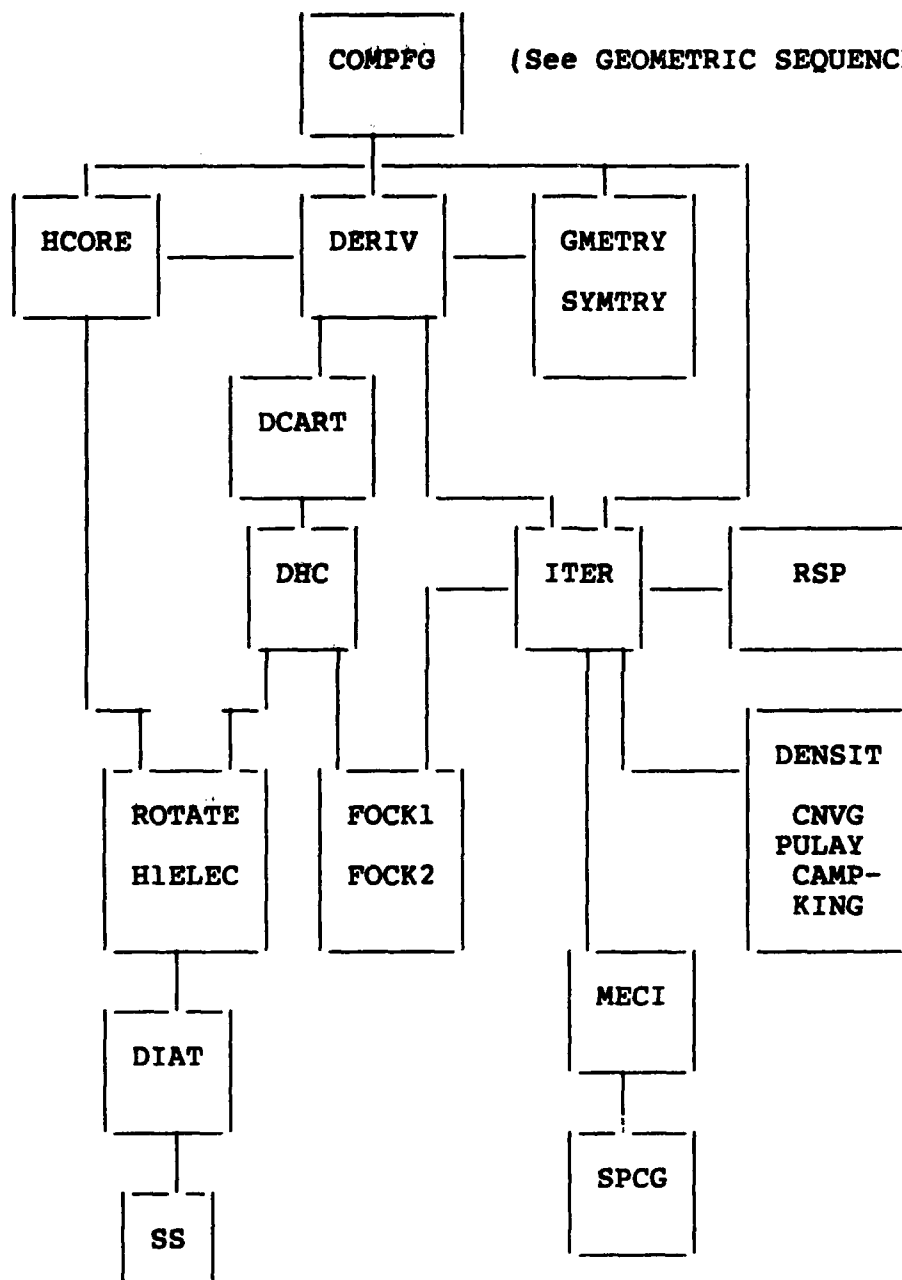
CHAPTER 7

PROGRAM

The logic within MOPAC is best understood by use of flow-diagrams.

There are two main sequences, geometric and electronic. These join only at one common subroutine COMPFG. It is possible, therefore, to understand the geometric or electronic sections in isolation, without having studied the other section.

7.1 MAIN GEOMETRIC SEQUENCE

```
                              ┌──────────┐
                              │   MAIN   │
                              └────┬─────┘
        ┌──────────┬──────────────┼──────────────┬─────────────┐
        │     ┌────┴───┐     ┌─────┴────┐   ┌──────┴───┐   ┌────┴────┐
        │     │  IRC   │     │  FORCE   │   │  REACT1  │   │  PATHS  │
        │     │  and   │     └────┬─────┘   └────┬─────┘   └─────────┘
        │     │  DRC   │          │              │
        │     └────┬───┘          │              │
   ┌────┴────┐     │        ┌─────┴────┐   ┌──────┴───┐
   │ NLLSQ   │     │        │   FMAT   │   │  FLEPO   │
   │  and    │     │        └────┬─────┘   └────┬─────┘
   │ POWSQ   │     │             │              │
   └────┬────┘     │             │        ┌─────┴────┐
 ┌──────┴───┐      │             │        │  LINMIN  │
 │ SEARCH   │      │             │        └──────────┘
 │   or     │      │             │
 │ LOCMIN   │      │             │
 └────┬─────┘      │             │
      └────────────┴─────────────┤
                           ┌─────┴────┐
                           │  COMPFG  │   (See ELECTRONIC SEQUENCE)
                           └──────────┘
```

7.2 MAIN ELECTRONIC FLOW

```
                        ┌─────────┐
                        │ COMPFG  │   (See GEOMETRIC SEQUENCE)
                        └─────────┘
         ┌──────────────────┼──────────────────┐
    ┌─────────┐        ┌─────────┐        ┌─────────┐
    │  HCORE  │ ────── │  DERIV  │ ────── │ GMETRY  │
    │         │        │         │        │         │
    └─────────┘        └─────────┘        │ SYMTRY  │
         │                  │             │         │
         │             ┌─────────┐        └─────────┘
         │             │  DCART  │             │
         │             │         │             │
         │             └─────────┘             │
         │                  │                  │
         │             ┌─────────┐   ┌─────────┐    ┌─────────┐
         │             │   DHC   │───│  ITER   │────│   RSP   │
         │             │         │   │         │    │         │
         │             └─────────┘   └─────────┘    └─────────┘
         │                  │             │
         │                  │             │         ┌─────────┐
    ┌─────────┐        ┌─────────┐        │         │ DENSIT  │
    │ ROTATE  │        │ FOCK1   │        │         │         │
    │         │        │         │        │         │  CNVG   │
    │ H1ELEC  │        │ FOCK2   │        │         │ PULAY   │
    │         │        │         │        │         │ CAMP-   │
    └─────────┘        └─────────┘        │         │ KING    │
         │                               │         └─────────┘
    ┌─────────┐                     ┌─────────┐
    │  DIAT   │                     │  MECI   │
    │         │                     │         │
    └─────────┘                     └─────────┘
         │                               │
    ┌─────────┐                     ┌─────────┐
    │   SS    │                     │  SPCG   │
    │         │                     │         │
    └─────────┘                     └─────────┘
```

7.3 CONTROL WITHIN MOPAC

Almost all the control information is passed via the single datum "KEYWRD", a string of 80 characters, which is read in at the start of the job.

Each subroutine is made independent, as far as possible, even at the expense of extra code or calculation. Thus, for example, the SCF criterion is set in subroutine ITER, and nowhere else. Similarly subroutine DERIV has exclusive control of the step size in the finite-difference calculation of the energy derivatives. If the default values are to be reset, then the new value is supplied in KEYWRD, and extracted via INDEX and READA. The flow of control is decided by the presence of various keywords in KEYWRD.

When a subroutine is called, it assumes that all data required for its operation are available in either common blocks or arguments. Normally no check is made as to the validity of the data received. All data are "owned" by one, and only one, subroutine. By ownership is implied the permission and ability to change the data. Thus MOLDAT "owns" the number of atomic orbitals, in that it calculates this number, and stores it in the variable NORBS. Many subroutines use NORBS, but none of them is allowed to change it. For obvious reasons no exceptions should be made to this rule. To illustrate the usefulness of this convention, consider the eigenvectors, C and CBETA. These are owned by ITER. Before ITER is called, C and CBETA are not calculated, after ITER has been called C and CBETA are known, so any subroutine which needs to use the eigenvectors can do so in the certain knowledge that they exist.

Any variables which are only used within a subroutine are not passed outside the subroutine unless an overriding reason exists. This is found in PULAY and CNVG, among others where arrays used to hold spin-dependent data are used, and these cannot conveniently be defined within the subroutines. In these examples, the relevant arrays are "owned" by ITER.

A general subroutine, of which ITER is a good example, handles three kinds of data: First, data which the subroutine is going to work on, for example the one and two electron matrices; second, data necessary to manipulate the first set of data, such as the number of atomic orbitals; third, the calculated quantities, here the electronic energy, and the density and Fock matrices.

Reference data are entered into a subroutine by way of the common blocks. This is to emphasize their peripheral role. Thus the number of orbitals, while essential to ITER, is not central to the task it has to perform, and is passed through a common block.

Data the subroutine is going to work on are passed via the argument list. Thus the one and two electron matrices, which are the main reason for ITER's existence, are entered as two of the four arguments. As ITER does not own these matrices it can use them but may not change their contents. The other arguments are EE, the electronic energy, and FULSCF, a logical. EE is owned by ITER even though it first appears before ITER is called. FULSCF, on the other hand, is not owned by ITER,

and is used, but not changed.

Sometimes common block data should more correctly appear in an argument list. This is usually not done in order to prevent obscuring the main role the subroutine has to perform. Thus ITER calculates the density and Fock matrices, but these are not represented in the argument list as the calling subroutine never needs to know them; instead, they are stored in common.

SUBROUTINE GMETRY: Description for programmers.

GMETRY has two arguments, GEO and COORD. On input GEO contains either (a) internal coordinates or (b) cartesian coordinates. On exit COORD contains the cartesian coordinates.

The normal mode of usage is to supply the internal coordinates, in which case the connectivity relations are found in common block GEOKST.

If the contents of NA(1) is zero, as required for any normal system, then the normal internal to cartesian conversion is carried out.

If the contents of NA(1) is 99, then the coordinates found in GEO are assumed to be cartesian, and no conversion is made. This is the situation in a FORCE calculation.

A further option exists within the internal to cartesian conversion. If STEP, stored in common block REACTN, is non-zero, then a reaction path is assumed, and the internal coordinates are adjusted radially in order that the "distance" in internal coordinate space from the geometry specified in GEO is STEPP away from the geometry stored in GEOA, stored in REACTN.

During the internal to cartesian conversion, the angle between the three atoms used in defining a fourth atom is checked to ensure that it is not near to 0 or 180 degrees. If it is near to these angles, then there is a high probability that a faulty geometry will be generated and to prevent this the calculation is stopped and an error message printed.

NOTE 1: If the angle is exactly 0 or 180 degrees, then the calculation is not terminated: This is the normal situation in a high-symmetry molecule such as propyne.

NOTE 2: The check is only made if the fourth atom has a bond angle which is not zero or 180 degrees.

CHAPTER 8

ERROR MESSAGES PRODUCED BY MOPAC

MOPAC produces several hundred messages, all of which are intended to be self-explanatory. However, when an error occurs it is useful to have more information than is given in the standard messages.

The following alphabetical list gives more complete definitions of the messages printed.

AN UNOPTIMIZABLE GEOMETRIC PARAMETER....

When internal coordinates are supplied, six coordinates cannot be optimized. These are the three coordinates of atom 1, the angle and dihedral on atom 2 and the dihedral on atom 3. An attempt has been made to optimize one of these. This is usually indicative of a typographic error, but might simply be an oversight. Either way, the error will be corrected and the calculation will not be stopped here.

ATOM NUMBER nn IS ILLDEFINED

The rules for definition of atom connectivity are:

1. Atom 2 must be connected to atom 1 (default – no override)

2. Atom 3 must be connected to atom 1 or 2, and make an angle with 2 or 1.

3. All other atoms must be defined in terms of already-defined atoms: these atoms must all be different. Thus atom 9 might be connected to atom 5, make an angle with atom 6, and have a dihedral with atom 7. If the dihedral was with atom 5, then the geometry definition would be faulty.

If any of these rules is broken, a fatal error message is printed, and the calculation stopped.

ATOMIC NUMBER nn IS NOT AVAILABLE ...

An element has been used for which parameters are not available. Only if a typographic error has been made can this be rectified. This check is not exhaustive, in that even if the elements are acceptable there are some combinations of elements within MINDO/3 that are not allowed. This is a fatal error message.

ATOMIC NUMBER OF nn ?

An atom has been specified with a negative or zero atomic number. This is normally caused by forgetting to specify an atomic number or symbol. This is a fatal error message.

ATOMS   nn AND nn ARE SEPARATED BY nn.nnnn ANGSTROMS.

Two genuine atoms (not dummies) are separated by a very small distance. This can occur when a complicated geometry is being optimized, in which case the user may wish to continue. This can be done by using the keyword GEO-OK. More often, however, this message indicates a mistake, and the calculation is, by default, stopped.

ATTEMPT TO GO DOWNHILL IS UNSUCCESSFUL...

A quite rare message, produced by Bartel's gradient norm minimization. Bartel's method attempts to minimize the gradient norm by searching the gradient space for a minimum. Apparently a minimum has been found, but not recognized as such. The program has searched in all $(3N-6)$ directions, and found no way down, but the criteria for a minimum have not been satisfied. No advice is available for getting round this error.

BOTH SYSTEMS ARE ON THE SAME SIDE...

A non-fatal message, but still cause for concern. During a SADDLE calculation the two geometries involved are on opposite sides of the transition state. This situation is verified at every point by calculating the cosine of the angle between the two gradient vectors. For as long as it is negative, then the two geometries are on opposite sides of the T/S. If, however, the cosine becomes positive, then the assumption is made that one moiety has fallen over the T/S and is now below the other geometry. That is, it is now further from the T/S than the other, temporarily fixed, geometry. To correct this, identify geometries corresponding to points on each side of the T/S. (Two geometries on the output separated by the message "SWAPPING...") and make up a new data-file using these geometries. This corresponds to points on the reaction path near to the T/S. Run a new job using these two geometries, but with BAR set to a third or a quarter of its original value, e.g. BAR=0.05. This normally allows the T/S to be located.

C.I. NOT ALLOWED WITH UHF

There is no UHF configuration interaction calculation in MOPAC. Either remove the keyword that implies C.I. or the word UHF.

CALCULATION ABANDONED AT THIS POINT

A particularly annoying message! In order to define an atom's position, the three atoms used in the connectivity table must not accidentally fall into a straight line. This can happen during a geometry optimization or gradient minimization. If they do, and if the angle made by the atom being defined is not zero or 180 degrees, then its position becomes ill-defined. This is not desirable, and the calculation will stop in order to allow corrective action to be taken. Note that if the three atoms are in an exactly straight line, this message will not be triggered. The good news is that the criterion used to trigger this message was set too coarsely. The criterion has been tightened so that this message now does not often appear. Geometric integrity does not appear to be compromized.

CARTESIAN COORDINATES READ IN, AND CALCULATION...

If cartesian coordinates are read in, but the calculation is to be carried out using internal coordinates, then either all possible geometric variables must be optimized, or none can be optimized. If only some are marked for optimization then ambiguity exists. For example, if the "X" coordinate of atom 6 is marked for optimization, but the "Y" is not, then when the conversion to internal coordinates takes place, the first coordinate becomes a bond-length, and the second an angle. These bear no relationship to the "X" or "Y" coordinates. This is a fatal error.

CARTESIAN COORDINATES READ IN, AND SYMMETRY...

If cartesian coordinates are read in, but the calculation is to be carried out using internal coordinates, then any symmetry relationships between the cartesian coordinates will not be reflected in the internal coordinates. For example, if the "Y" coordinates of atoms 5 and 6 are equal, it does not follow that the internal coordinate angles these atoms make are equal. This is a fatal error.

ELEMENT NOT FOUND

When an external file is used to redefine MNDO or AM1 parameters, the chemical symbols used must correspond to known elements. Any that do not will trigger this fatal message.

ERROR DURING READ AT ATOM NUMBER ....

Something is wrong with the geometry data. In order to help find the error, the geometry already read in is printed. The error lies either on the last line of the geometry printed, or on the next (unprinted) line. This is a fatal error.

FAILED IN SEARCH, SEARCH CONTINUING

Not a fatal error. The McIver-Komornicki gradient minimization involves use of a line-search to find the lowest gradient. This message is merely advice. However, if SIGMA takes a long time, consider doing something else, such as using NLLSQ, or refining the geometry a bit before resubmitting it to SIGMA.

<<<<----**** FAILED TO ACHIEVE SCF. ****---->>>>

The SCF calculation failed to go to completion; an unwanted and depressing message that unfortunately appears every so often.

To date three unconditional convergers have appeared in the literature: the SHIFT technique, Pulay's method, and the Camp-King converger. It would not be fair to the authors to condemn their methods. In MOPAC all sorts of weird and wonderful systems are calculated, systems the authors of the convergers never dreamed of. MOPAC uses a combination of all three convergers at times. Normally only a quadratic damper is used.

If this message appears, suspect first that the calculation might be faulty, then, if you feel confident, use PL to monitor a single SCF. Based on the SCF results either increase the number of allowed iterations, default: 200, or use PULAY, or Camp-King, or a mixture.

If nothing works, then consider slackening the SCF criterion. This will allow heats of formation to be calculated with reasonable precision, but the gradients are likely to be imprecise.

GEOMETRY TOO UNSTABLE FOR EXTRAPOLATION..

In a reaction path calculation the initial geometry for a point is calculated by quadratic extrapolation using the previous three points.

If a quadratic fit is likely to lead to an inferior geometry, then the geometry of the last point calculated will be used. The total effect is to slow down the calculation, but no user action is recommended.

** GRADIENT IS TOO LARGE TO ALLOW...

Before a FORCE calculation can be performed the gradient norm must be so small that the third and higher order components of energy in the force field are negligible. If, in the system under examination, the gradient norm is too large, the gradient norm will first be reduced using FLEPO, unless LET has been specified. In some cases the FORCE calculation may be run only to decide if a state is a ground state or a transition state, in which case the results have only two interpretations. Under these circumstances, LET may be warranted.

GRADIENT IS VERY LARGE...

In a calculation of the thermodynamic properties of the system, if the rotation and translation vibrations are non-zero, as would be the case if the gradient norm was significant, then these "vibrations" would interfere with the low-lying genuine vibrations.   The criteria for THERMO are much more stringent than for a vibrational frequency calculation, as it is the lowest few genuine vibrations that determine the internal vibrational energy, entropy, etc.

ILLEGAL ATOMIC NUMBER

An element has been specified by an atomic number which is not in the range 1 to 107.   Check the data: the first datum on one of the lines is faulty. Most likely line 4 is faulty.

IMPOSSIBLE NUMBER OF OPEN SHELL ELECTRONS

The keyword OPEN(n1,n2) has been used, but for an even-electron system n1 was specified as odd or for an odd-electron system n1 was specified as even. Either way, there is a conflict which the user must resolve.

IMPOSSIBLE OPTION REQUESTED

A general catch-all. This message will be printed if two incompatible options are used, such as both MINDO/3 and AM1 being specified. Check the keywords, and resolve the conflict.

INTERNAL COORDINATES READ IN, AND CALCULATION...

If internal coordinates are read in, but the calculation is to be carried out using cartesian coordinates, then either all possible geometric variables must be optimized, or none can be optimized. If only some are marked for optimization, then ambiguity exists. For example, if the bond-length of atom 6 is marked for optimization, but the angle is not, then when the conversion to cartesian coordinates takes place, the first coordinate becomes the "X" coordinate and the second the "Y" coordinate. These bear no relationship to the bond length or angle. This is a fatal error.

INTERNAL COORDINATES READ IN, AND SYMMETRY...

If internal coordinates are read in, but the calculation is to be carried out using cartesian coordinates, then any symmetry relationships between the internal coordinates will not be reflected in the cartesian coordinates. For example, if the bond-lengths of atoms 5 and 6 are equal, it does not follow that these atoms have equal values for their "X" coordinates. This is a fatal error.

JOB STOPPED BY OPERATOR

Any MOPAC calculation, for which the SHUTDOWN command works, can be stopped by a user who issues the command "$SHUT <filename>, from the directory which contains <filename>.DAT

MOPAC will then stop the calculation at the first convenient point, usually after the current cycle has finished. A restart file will be written and the job ended. The message will be printed as soon as it is detected, which would be the next time the timer routine is accessed.

**** MAX. NUMBER OF ATOMS ALLOWED:....

At compile time the maximum sizes of the arrays in MOPAC are fixed. The system being run exceeds the maximum number of atoms allowed. To rectify this, modify the file DIMSIZES.DAT to increase the number of heavy and light atoms allowed. If DIMSIZES.DAT is altered, then the whole of MOPAC should be re-compiled and re-linked.

**** MAX. NUMBER OF ORBITALS:....

к
          could be used. If an attempt is made to modify any other parameter then an error message is printed, and the calculation terminated.

- 125 -

NUMBER OF PARTICLES, nn GREATER THAN...

When user-defined microstates are not used, the MECI will calculate all possible microstates that satisfy the space and spin constraints imposed. This is done in PERM, which permutes N electrons in M levels. If N is greater than M, then no possible permutation is valid. This is not a fatal error - the program will continue to run, but no C.I. will be done.

NUMBER OF PERMUTATIONS TOO GREAT, LIMIT 60

The number of permutations of alpha or beta microstates is limited to 60. Thus if 3 alpha electrons are permuted among 5 M.O.'s, that will generate 10 = 5!/(3!*2!) alpha microstates, which is an allowed number. However if 4 alpha electrons are permuted among 8 M.O.'s, then 70 alpha microstates result and the arrays defined will be insufficient. Note that 60 alpha and 60 beta microstates will permit 3600 microstates in all, which should be more than sufficient for most purposes. (An exception would be for excited radical icosohedral systems.)

SYMMETRY SPECIFIED, BUT CANNOT BE USED IN DRC

This is self explanatory. The DRC requires all geometric constraints to be lifted. Any symmetry constraints will first be applied, to symmetrize the geometry, and then removed to allow the calculation to proceed.

SYSTEM DOES NOT APPEAR TO BE OPTIMIZABLE

This is a gradient norm minimization message. These routines will only work if the nearest minimum to the supplied geometry in gradient-norm space is a transition state or a ground state. Gradient norm space can be visualized as the space of the scalar of the derivative of the energy space with respect to geometry. To a first approximation, there are twice as many minima in gradient norm space as there are in energy space.

It is unlikely that there exists any simple way to refine a geometry that results in this message. While it is appreciated that a large amount of effort has probably already been expended in getting to this point, users should steel themselves to writing off the whole geometry. It is not recommended that a minor change be made to the geometry and the job re-submitted.

Try using SIGMA instead of POWSQ.

TEMPERATURE RANGE STARTS TOO LOW,...

The thermodynamics calculation assumes that the statistical summations can be replaced by integrals. This assumption is only valid above 100K, so the lower temperature bound is set to 100, and the calculation continued.

THERE IS A RISK OF INFINITE LOOPING...

The SCF criterion has been reset by the user, and the new value is so small that the SCF test may never be satisfied. This is a case of user beware!

THIS MESSAGE SHOULD NEVER APPEAR, CONSULT A PROGRAMMER!

This message should never appear; a fault has been introduced into MOPAC, most probably as a result of a programming error. If this message appears in the vanilla version of MOPAC (A version ending in 0), please contact JJPS as I would be most interested in how this was achieved.

THREE ATOMS BEING USED TO DEFINE....

If the cartesian coordinates of an atom depend on the dihedral angle it makes with three other atoms, and those three atoms fall in an almost straight line, then a small change in the cartesian coordinates of one of those three atoms can cause a large change in its position. This is a potential source of trouble, and the data should be changed to make the geometric specification of the atom in question less ambiguous.

This message can appear at any time, particularly in reaction path and saddle-point calculations.

An exception to this rule is if the three atoms fall into an exactly straight line. For example, if, in propyne, the hydrogens are defined in terms of the three carbon atoms, then no error will be flagged. In such a system the three atoms in the straight line must not have the angle between them optimized, as the finite step in the derivative calculation would displace one atom off the straight line and the error-trap would take effect.

Correction involves re-defining the connectivity. LET and GEO-OK will not allow the calculation to proceed.

- - - - - - - TIME UP - - - - - - -

   The time defined on the keywords line or 3,600 seconds, if no time
was specified, is likely to be exceeded if another cycle of calculation
were to be performed. A controlled termination of the run would follow
this message. The job may terminate earlier than expected: this is
ordinarily due to one of the recently completed cycles taking unusually
long, and the safety margin has been increased to allow for the
possibility that the next cycle might also run for much longer than
expected.

                 TRIPLET SPECIFIED WITH ODD NUMBER OF ELECTRONS.

   If TRIPLET has been specified the number of electrons must be even.
Check the charge on the system, the empirical formula, and whether
TRIPLET was intended.

          """"""""""""""""UNABLE TO ACHIEVE SELF-CONSISTENCY

   See the error-message:

   <<<<----**** FAILED TO ACHIEVE SCF. ****---->>>>

                     UNDEFINED SYMMETRY FUNCTION USED

   Symmetry operations are restricted to those defined, i.e.   in the
range 1-18. Any other symmetry operations will trip this fatal message.

                       UNRECOGNIZED ELEMENT NAME

   In the geometric specification a chemical symbol which does not
correspond to any known element has been used. The error lies in the
first datum on a line of geometric data.

**** WARNING ****

Don't pay too much attention to this message. Thermodynamics calculations require a higher precision than vibrational frequency calculations. In particular, the gradient norm should be very small. However, it is frequently not practical to reduce the gradient norm further, and to date no-one has determined just how slack the gradient criterion can be before unacceptable errors appear in the thermodynamic quantities. The 0.4 gradient norm is only a suggestion.

WARNING: INTERNAL COORDINATES...

Triatomics are, by definition, defined in terms of internal coordinates. This warning is only a reminder. For diatomics, cartesian and internal coordinates are the same. For tetra-atomics and higher, the presence or absence of a connectivity table distinguishes internal and cartesian coordinates, but for triatomics there is an ambiguity. To resolve this, cartesian coordinates are not allowed for the data input for triatomics.

CHAPTER 9

CRITERIA

MOPAC uses various criteria which control the precision of its stages. These criteria are chosen as the best compromise between speed and acceptable errors in the results. The user can override the default settings by use of keywords; however, care should be exercised as increasing a criterion can introduce the potential for infinite loops, and decreasing a criterion can result in unacceptably imprecise results. These are usually characterized by 'noise' in a reaction path, or large values for the trivial vibrations in a force calculation.

9.1 SCF CRITERION

Name: SCFCRT.
Defined in   ITER.
Default value   0.00001 kcal/mole
Basic Test   Change in energy in kcal/mole on successive
                iterations is less than SCFCRT.

Exceptions: If PRECISE is specified,   SCFCRT=0.0000001
                If non-variational wavefunction SCFCRT=0.0000001
                If a polarization calculation
                or gradient minimization   SCFCRT=0.0000001
                If SCFCRT=n.nnn is specified   SCFCRT=n.nnn

Secondary tests: (1) Change in density matrix elements on two
                     successive iterations must be less than 0.001
                 (2) Change in energy in eV on three successive
                     iterations must be less than 10 x SCFCRT.

9.2 GEOMETRIC OPTIMIZATION CRITERIA

Name: TOLERX   "Test on X Satisfied"
Defined in   FLEPO
Default value   0.0001 Angstroms
Basic Test   The projected change in geometry is less than
                 TOLERX Angstroms.

Exceptions   If PRECISE is specified, TOLERX= 0.00001
                 If GNORM is specified, the TOLERX test is not used.

Name: EYEAD   "Herbert's Test Satisfied"
Defined in   FLEPO
Default value   0.001
Basic Test   The projected decrease in energy is less than
                 EYEAD Kcals/mole.

Exceptions   If PRECISE is specified, EYEAD=0.00001
                 If GNORM is specified, the EYEAD test is not used.

Name: TOLERG   "Test on Gradient Satisfied"
Defined in   FLEPO
Default value   1.0
Basic Test   The gradient norm in Kcals/mole/Angstrom is less
                 than TOLERG multiplied by the square root of the
                 number of coordinates to be optimized.

Exceptions   If PRECISE is specified, TOLERG=0.01
                 If GNORM=n.nnn is specified, TOLERG=n.nnn divided
                 by the square root of the number of coordinates
                 to be optimized, and the secondary tests are not
                 done.

                 If a SADDLE calculation, TOLERG is made a function
                 of the last gradient norm.
Name: TOLERF   "Heat of Formation Test Satisfied"
Defined in   FLEPO
Default value   0.002 Kcal/mole
Basic Test   The calculated heats of formation on two successive
                 cycles differ by less than TOLERF.

Exceptions   If PRECISE is specified, TOLERF=0.00004
                 If GNORM is specified, the TOLERF test is not used.

Secondary Tests For the TOLERG, TOLERF, and TOLERX tests, a
                 second test in which no individual component of the
                 gradient should be larger than TOLERG must be
                 satisfied.

Other Tests   If, after the TOLERG, TOLERF, or TOLERX test has been
                 satisfied three consecutive times the heat of
                 formation has dropped by less than 0.3Kcal/mole, then
                 the optimization is stopped.

Exceptions   If GNORM is specified, then this test is not performed.

Name:	TOL2
Defined in	POWSQ
Default value	0.4
Basic Test	The absolute value of the largest component of the gradient is less than TOL2

Exceptions	If PRECISE is specified, TOL2=0.01

Name:	TOLS1
Defined in	NLLSQ
Default Value	0.000 000 000 001
Basic Test	The square of the ratio of the projected change in the geometry to the actual geometry is less than TOLS1.

Name:	\<none\>
Defined in	NLLSQ
Default Value	0.2
Basic Test	Every component of the gradient is less than 0.2.

CHAPTER 10

DEBUGGING

There are three potential sources of difficulty in using MOPAC, each of which requires special attention. There can be problems with data, due to errors in the data, or MOPAC may be called upon to do calculations for which it was not designed. There are intrinsic errors in MOPAC which extensive testing has not yet revealed, but which a user's novel calculation uncovers. Finally there can be bugs introduced by the user modifying MOPAC, either to make it compatible with the host computer, or to implement local features.

For whatever reason, the user may need to have access to more information than the normal keywords can provide, and a second set, specifically for debugging, is provided. These keywords give information about the working of individual subroutines, and do not affect the course of the calculation.

10.1 DEBUGGING KEYWORDS

FULL LIST OF KEYWORDS FOR DEBUGGING SUBROUTINES
INFORMATION PRINTED

1ELEC	the one-electron matrix.	Note 1
COMPFG	Heat of Formation.	
DCART	Cartesian derivatives.	
DEBUG		Note 2
DEBUGPULAY	Pulay matrix, vector, and error-function.	Note 3
DENSITY	Every density matrix.	Note 1
DERIV	All gradients, and other data in DERIV.	
DFORCE	Print Force Matrix.	
EIGS	All eigenvalues.	
FLEPO	Details of BFGS minimization.	
FMAT		
FOCK	Every Fock matrix	Note 1
HCORE	The one electron matrix, and two electron integrals.	
ITER	Values of variables and constants in ITER.	
MOLDAT	Molecular data, number of orbitals, "U" values, etc.	
MECI	C.I. matrices, M.O. indices, etc.	
PL	Differences between density matrix elements	Note 4

	in ITER.	
LINMIN	Function values, step sizes at all points in the line minimization (LINMIN or SEARCH).	
TIMES	Times of stages within ITER.	
VECTORS	All eigenvectors on every iteration.	Note 1

NOTES

1. These keywords are activated by the keyword DEBUG. Thus if DEBUG and FOCK are both specified, every Fock matrix on every iteration will be printed.

2. DEBUG is not intended to increase the output, but does allow other keywords to have a special meaning.

3. PULAY is already a keyword, so DEBUGPULAY was an obvious alternative.

4. PL initiates the output of the value of the largest difference between any two density matrix elements on two consecutive iterations. This is very useful when investigating options for increasing the rate of convergence of the SCF calculation.

SUGGESTED PROCEDURE FOR LOCATING BUGS

Users are supplied with the source code for MOPAC, and, while the original code is fairly bug-free, after it has been modified there is a possibility that bugs may have been introduced. In these circumstances the author of the changes is obviously responsible for removing the offending bug, and the following ideas might prove useful in this context.

First of all, and most important, before any modifications are done a back-up copy of the standard MOPAC should be made. This will prove invaluable in pinpointing deviations from the standard working. This point cannot be over-emphasized - MAKE A BACK-UP BEFORE MODIFYING MOPAC!!!!

Clearly, a bug can occur almost anywhere, and a logical search sequence is necessary in order to minimize the time taken to locate it.

If possible, perform the debugging with a small molecule, in order to save time (debugging is, of necessity, time consuming) and to minimize output.

The two sets of subroutines in MOPAC, those involved with the electronics and those involved in the geometrics, are kept strictly separate, so the first question to be answered is which set contains the bug. If the heats of formation, derivatives, I.P.s, and charges, etc., are correct, the bug lies in the geometrics; if faulty, in the electronics.

Bug in the Electronics Subroutines.

Use formaldehyde for this test. The supplied data-file MNRSD1.DAT could be used as a template for this operation. Use keywords 1SCF, DEBUG, and any others necessary.

The main steps are:

(1) Check the starting one-electron matrix and two-electron integral string, using the keyword HCORE. It is normally sufficient to verify that the two hydrogen atoms are equivalent, and that the pi system involves only pz on oxygen and carbon. Note that numerical values are not checked, but only relative values.

If an error is found, use MOLDAT to verify the orbital character, etc.

If faulty the error lies in READ, GETGEO or MOLDAT.

Otherwise the error lies in HCORE, H1ELEC or ROTATE.

If the starting matrices are correct, go on to step (2).

(2) Check the density or Fock matrix on every iteration, with the words FOCK or DENSITY. Check the equivalence of the two hydrogen atoms, and the pi system, as in (1).

If an error is found, check the first Fock matrix. If faulty, the bug lies in ITER, probably in the Fock subroutines FOCK1 or FOCK2. or in the (guessed) density matrix (MOLDAT). An exception is in the UHF closed-shell calculation, where a small asymmetry is introduced to initiate the separation of the alpha and beta UHF wavefunctions.

If no error is found, check the second Fock matrix. If faulty, the error lies in the density matrix DENSIT, or the diagonalization RSP.

If the Fock matrix is acceptable, check all the Fock matrices. If the error starts in iterations 2 to 4, the error probably lies in CNVG, if after that, in PULAY, if used.

If SCF is achieved, and the heat of formation is faulty, check HELECT. If C.I. was used check MECI.

If the derivatives are faulty, use DCART to verify the cartesian derivatives. If these are faulty, check DCART and DHC. If they are correct, or not calculated, check the DERIV finite difference calculation.

If the geometric calculation is faulty, use FLEPO to monitor the optimization, DERIV may also be useful here.

For the FORCE calculation, DCART or DERIV are useful for variationally optimized functions, COMPFG for non-variationally optimized functions.

For reaction paths, verify that FLEPO is working correctly; if so, then **PATHS** is faulty.

For saddle-point calculations, verify that FLEPO is working correctly; if so, then **REACT1** is faulty.

Keep in mind the fact that MOPAC is a large calculation, and while intended to be versatile, many combinations of options have not been tested. If a bug is found in the original code, please communicate details to the Academy, to Dr. James J. P. Stewart, Frank J. Seiler Research Laboratory, U.S. Air Force Academy, Colorado Springs, CO 80840-6528.

CHAPTER 11

INSTALLING MOPAC

Specific instructions for mounting MOPAC on other computers have been left out due to limitations of space in the Manual.

How to use MOPAC

MOPAC Restarts should be user transparent. If MOPAC does make any restart files, do not change them (It would be hard to do anyhow, as they're in machine code), as they will be used when you run a RESTART job. The main files that are produced are:

```
<filename>.OUT    Results
<filename>.ARC    Archive or summary
<filename>.RES    Restart
<filename>.DEN    Density matrix (in binary)
```

Size of MOPAC

The amount of storage required by MOPAC depends mainly on the number of heavy and light atoms. As it is useful for programmers to have an idea of how large various MOPACs are, the following data are presented as a guide.

Sizes of various MOPAC Version 5.00 executables in which the number of heavy atoms is equal to the number of light atoms, assembled on a VAX computer are:

No. of heavy atoms	Size of Executable (bytes)
10	1,653,760
20	3,442,688
30	6,356,480
40	10,400,768
50	15,572,480
60	21,872,128
100	58,361,856
200	228,602,880
300	511,723,008
400	907,723,264

The size of any given MOPAC executable may be estimated from

$$Size = 993868 + N*9571 + N*N*5640 \text{ bytes}$$

The size of MOPAC executables will vary from machine to machine, due to the different sizes of the code. For a VAX, this amounts to approximately 0.1Mb. Most machines use a 64 bit or 8 byte double precision real number, so the 9571 and 5640 multipliers should apply to them. For large jobs, 0.1Mb is negligable, therefore the above expression should be applicable to most computers.

No. of lines in program in Version 5.00 = 22,084
                                        = 17,718 code + 4,366 comment

APPENDIX A

FORTRAN FILES

NAMES OF FORTRAN-77 FILES

AABABC	ANALYT	ANAVIB	AXIS	BLOCK
BONDS	CALPAR	CAPCOR	CHRGE	CNVG
COMPFG	DATIN	DCART	DELMOL	DELRI
DENROT	DENSIT	DEPVAR	DERIV	DERS
DFPSAV	DIAG	DIAT	DIAT2	DIPOLE
DIPINV	DOT	DRC	DRCOUT	ENPART
EXCHNG	FFHPOL	FLEPO	FMAT	FOCK1
FOCK2	FOCK2D	FORCE	FORSAV	FRAME
FREQCY	GEOUT	GETGEO	GETSYM	GMETRY
GOVER	GRID	H1ELEC	HADDON	HCORE
HELECT	HQRII	IJKL	INTERP	ITER
LOCAL	LOCMIN	MAMULT	MATOUT	MECI
MNDO	MOLDAT	MOLVAL	MULLIK	MULT
NLLSQ	NUCHAR	OSINV	PARSAV	PATHS
PERM	POLAR	POWSAV	POWSQ	PRTDRC
PULAY	QUADR	REACT1	READ	READA
REFER	REPP	ROTAT	ROTATE	RSP
SEARCH	SECOND	SETUPG	SOLROT	SPCG
SWAP	SYMTRY	THERMO	UPDATE	VECPRT
WRITE	WRTKEY	XYZINT		

APPENDIX B

SUBROUTINE CALLS IN MOPAC

A list of the program segments which call various subroutines.

Subroutine	Called by
AM1	MOPAC
ANALYT	DCART
ANAVIB	FORCE
AXIS	FORCE FRAME
BANGLE	XYZGEO XYZINT
BFN	SS
BONDS	WRITE
CALPAR	AM1
CHRGE	DRC FMAT FOCK1 POWSQ WRITE
CNVG	ITER
COE	DENROT DIAT
COMPFG	DRC FLEPO FMAT FORCE MOPAC NLLSQ POLAR POWSQ REACT1 SEARCH
DANG	DIHED
DATE	WRITE

Subroutine	Called by
DCART	DERIV
DELMOL	ANALYT
DELRI	ANALYT
DENSIT	ITER MULLIK
DEPVAR	HADDON
DERIV	COMPFG
DERS	ANALYT
DFPSAV	FLEPO PATHS
DHC	DCART
DIAG	ITER
DIAT	H1ELEC
DIAT2	DIAT
DIHED	XYZGEO XYZINT
DIPOLE	WRITE FMAT
DIPIND	POLAR
DRC	FORCE MOPAC
DRCOUT	PRTDRC
ENPART	WRITE
ENPART	WRITE
EPSETA	RSP DIAG
EXCHNG	SEARCH
FFHPOL	POLAR
FLEPO	FORCE GRID MOPAC PATHS REACT1
FMAT	FORCE
FOCK1	ITER
FOCK2	ITER

FOCK2D   DHC

Subroutine	Called by
FORCE	MOPAC
FORSAV	FMAT
FRAME	FMAT FORCE FREQCY
FREQCY	FORCE
GEOUT	GMETRY READ REACT1 WRITE
GETGEO	REACT1 READ
GETSYM	READ
GMETRY	COMPFG DERIV DRC FORCE MOLDAT MULLIK POLAR READ WRITE
GOVER	DIAT
GRID	MOPAC
H1ELEC	DHC HCORE
HADDON	SYMTRY
HCORE	COMPFG DERIV
HQRII	INTERP POWSQ
IJKL	MECI
INTERP	ITER
ITER	COMPFG DERIV
LOCAL	WRITE

Subroutine	Called by					
MAMULT	PULAY					
MATOUT	FMAT	FORCE	INTERP	ITER	LOCAL	MECI
MECI	ITER	WRITE				
MOLDAT	AM1	MOPAC				
MOLVAL	WRITE					
MULLIK	WRITE					
MULT	MULLIK					
NLLSQ	FORCE	MOPAC				
NUCHAR	GETSYM	READ				
OSINV	PULAY					
PARSAV	NLLSQ					
PATHS	MOPAC					
PERM	MECI					
POLAR	MOPAC					
POWSAV	POWSQ					
POWSQ	MOPAC					
PRTDRC	DRC					
PULAY	ITER					
QUADR	PRTDRC					
REACT1	MOPAC					
READ	MOPAC					
READA	WRTKEY	WRITE	THERMO	REACT1	PRTDRC	POWSQ
	NUCHAR	NLLSQ	MOLDAT	MATOUT	ITER	GRID
	GETGEO	FORCE	FMAT	FLEPO	DRC	DERIV
	AM1	DEPVAR				

Subroutine	Called by
REFER	MOLDAT
REPP	ROTATE
ROTAT	DELMOL
ROTATE	DHC HCORE SOLROT
RSP	AXIS FMAT FORCE FRAME ITER MECI MULLIK POLAR
SCHMIB	HQRII
SCHMIT	HQRII
SEARCH	POWSQ NLLSQ
SET	DIAT2
SETUPG	COMPFG
SOLROT	DHC HCORE
SPLINE	HQRII
SWAP	ITER
SYMTRY	COMPFG DERIV REACT1 READ WRITE
THERMO	FORCE
TQL2	RSP
TQLRAT	RSP
TRBAK3	RSP
TRED3	RSP
UPDATE	AM1
VECPRT	BONDS FORCE HCORE INTERP ITER MECI MOLDAT MULLIK POLAR POWSQ SOLROT HADDON WRITE
WRITE	FORCE ITER MOPAC PATHS REACT1
WRTKEY	READ
XYZGEO	XYZINT
XYZINT	DFPSAV DRC FORCE GEOUT GETGEO PARSAV POWSQV PRTDRC WRITE

A list of subroutines called by various segments (the inverse of the first list)

SUBROUTINE	CALLS
AM1	UPDATE MOLDAT CALPAR
ANALYT	DERS DELRI DELMOL
AXIS	RSP
BONDS	VECPRT
COMPFG	SETUPG SYMTRY GMETRY HCORE ITER DERIV
DCART	ANALYT DHC
DELMOL	ROTAT
DENROT	GMETRY COE
DERIV	SYMTRY GMETRY HCORE ITER DCART
DFPSAV	XYZINT
DHC	H1ELEC ROTATE SOLROT FOCK2D FOCK2D
DIAT	COE GOVER DIAT2
DIAT2	SET
DIHED	DANG
DRC	GMETRY COMPFG PRTDRC
FLEPO	DFPSAV COMPFG
FMAT	FORSAV COMPFG CHRGE FRAME RSP MATOUT
FOCK1	CHRGE
FORCE	COMPFG NLLSQ FLEPO WRITE GMETRY XYZINT AXIS FMAT VECPRT FRAME RSP MATOUT FREQCY DRC ANAVIB THERMO
FRAME	AXIS
FREQCY	FRAME RSP

SUBROUTINE	CALLS
GEOUT	XYZINT
GETGEO	XYZINT
GETSYM	NUCHAR
GMETRY	GEOUT
GRID	FLEPO
H1ELEC	DIAT
HADDON	DEPVAR
HCORE	H1ELEC ROTATE SOLROT VECPRT
INTERP	VECPRT HQRII MATOUT SCHMIT SCHMIB MATOUT SYSTEM SPLINE
ITER	VECPRT FOCK2 FOCK1 INTERP PULAY DIAG RSP MATOUT DENSIT CNVG WRITE
LOCAL	MATOUT
MECI	IJKL PERM VECPRT RSP MATOUT
MOLDAT	REFER GMETRY VECPRT
MULLIK	RSP GMETRY MULT DENSIT VECPRT
NLLSQ	COMPFG PARSAV GEOUT SEARCH
PARSAV	XYZINT
PATHS	DFPSAV FLEPO WRITE
POLAR	GMETRY COMPFG VECPRT RSP MATOUT DIPIND FFHPOL
POWSAV	XYZINT
POWSQ	POWSAV COMPFG VECPRT HQRII SEARCH
PRTDRC	CHRGE XYZINT QUADR DRCOUT
PULAY	MAMULT OSINV

SUBROUTINE	CALLS
REACT1	GETGEO SYMTRY GEOUT GMETRY FLEPO COMPFG WRITE
READ	GETGEO WRTKEY GETSYM SYMTRY NUCHAR GEOUT GMETRY
ROTATE	REPP
RSP	EPSETA TRED3 TQLRAT TQL2 TRBAK3
SEARCH	COMPFG
SOLROT	ROTATE VECPRT
SS	BFN
SYMTRY	HADDON
WRITE	DATE GEOUT DERIV SYMTRY GMETRY VECPRT MATOUT CHRGE DENROT MOLVAL BONDS LOCAL ENPART MULLIK XYZINT
XYZGEO	BANGLE DIHED
XYZINT	DIHED BANGLE XYZGEO

APPENDIX C

DESCRIPTION OF SUBROUTINES IN MOPAC

AABABC   Utility: Calculates the configuration interaction matrix
          element between two configurations differing by exactly
          one alpha M.O. Called by MECI only.

AABACD   Utility: Calculates the configuration interaction matrix
          element between two configurations differing by exactly
          two alpha M.O.'s. Called by MECI only.

AABBCD   Utility: Calculates the configuration interaction matrix
          element between two configurations differing by exactly
          two M.O.'s; one configuration has alpha M.O. "A" and beta
          M.O. "C" while the other configuration has alpha M.O. "B"
          and beta M.O. "D". Called by MECI only.

AINTGS   Utility: Within the overlap integrals, calculates the
          A-integrals. Dedicated to function SS within DIAT.

AM1   Utility: Reads in external parameters for use within
          MOPAC. Originally used for the testing of new parameters,
          AM1 is now a general purpose reader for parameters.
          Invoked by the keyword EXTERNAL.

ANALYT   Main Sequence: Calculates the analytical derivatives
          of the energy with respect to cartesian coordinates for all
          atoms. Use only if the mantissa is short (less than 52 bits)
          or out of interest. Should not be used for routine work
          on a VAX.

ANAVIB   Utility: Gives a brief interpretation of the modes of
          vibration of the molecule. The principal pairs of atoms
          involved in each vibration are identified, and the mode
          of motion (tangential or radial) is output.

AXIS   Utility: Works out the three principal moments of inertia
          of a molecule. If the system is linear, one moment of inertia
          is zero. Prints moments in units of $cm^{**}(-1)$ and
          $10^{**}(-40)$ gram-cm-cm.

BABBBC   Utility: Calculates the configuration interaction matrix

element between two configurations differing by exactly
one beta M.O. Called by MECI only.

BABBCD   Utility: Calculates the configuration interaction matrix
          element between two configurations differing by exactly
          two beta M.O.'s. Called by MECI only.

BANGLE   Utility: Given a set of coordinates, BANGLE will calculate the
          angle between any three atoms.

BFN   Utility: Calculates the B-functions in the Slater overlap.

BINTGS   Utility: Calculates the B-functions in the Slater overlap.

BONDS   Utility: Evaluates and prints the valencies of atoms and
          bond-orders between atoms. Main argument: density matrix.
          No results are passed to the calculation, and no data
          are changed. Called by WRITE only.

CALPAR   Utility: When external parameters are read in via EXTERNAL=,
          the derived parameters are worked out using CALPAR. Note that
          all derived parameters are calculated for all parameterized
          elements at the same time.

CAPCOR   Utility: Capping atoms, of type Cb, should not contribute to
          the energy of a system. CAPCOR calculates the energy
          contribution due to the Cb and subtracts it from the
          electronic energy.

CHRGE   Utility: Calculates the total number of valence electrons
          on each atom. Main arguments: density matrix, array of
          atom charges (empty on input). Called by ITER only.

CNVG   Utility: Used in SCF cycle. CNVG does a three-point
          interpolation of the last three density matrices.
          Arguments: Last three density matrices, Number of iterations,
          measure of self-consistency (empty on input). Called by ITER
          only.

COE   Utility: Within the general overlap routine COE calculates
          the angular coefficients for the s, p and d real atomic
          orbitals given the axis and returns the rotation matrix.

COMPFG   Main Sequence: Evaluates the total heat of formation of the
          supplied geometry, and the derivatives, if requested. This
          is the nodal point connecting the electronic and geometric
          parts of the program. Main arguments: on input: geometry,
          on output: heat of formation, gradients.

DCART   Utility: Called by DERIV, DCART sets up a list of cartesian
          derivatives of the energy W.R.T. coordinates which DERIV can
          then use to construct the internal coordinate derivatives.

DELMOL   Utility: Part of analytical derivates. Two-electron.

DELRI   Utility: Part of analytical derivates. Two-electron.

DENROT   Utility: Converts the ordinary density matrix into
            a condensed density matrix over basis functions s (sigma),
            p (sigma) and p (pi), i.e., three basis functions. Useful
            in hybridization studies. Has capability to handling "d"
            functions, if present.

DENSIT   Utility: Constructs the Coulson electron density matrix from
            the eigenvectors. Main arguments: Eigenvectors, No. of singly
            and doubly occupied levels, density matrix (empty on input)
            Called by ITER.

DEPVAR   Utility: A symmetry-defined "bond length" is related to
            another bond length by a multiple. This special symmetry
            function is intended for use in Cluster calculations.
            Called by HADDON.

DERIV   Main Sequence: Calculates the derivatives of the energy with
            respect to the geometric variables. This is done either by
            using initially cartesian derivatives (normal mode) or by full
            SCF calculations (half-electron and C.I. mode). Arguments: on
            input: geometry, on output: derivatives. Called by COMPFG.

DERS   Utility: Called by ANALYT, DERS calculates the analytical
            derivatives of the overlap matrix within the molecular frame.

DFPSAV   Utility: Saves and restores data used by the
            BFGS geometry optimization. Main arguments:
            parameters being optimized, gradients of parameters, last heat
            of formation, integer and real control data. Called by FLEPO.

DHC   Utility: Called by DCART and calculates the energy of a pair
            of atoms using the SCF density matrix. Used in the finite
            difference derivatve calculation.

DIAG   Utility: Rapid pseudo-diagonalization. Given a set of vectors
            which almost block-diagonalize a secular determinant, DIAG
            modifies the vectors so that the block-diagonalization is more
            exact. Main arguments: Old vectors, Secular Determinant,
            New vectors (on output). Called by ITER.

DIAGI   Utility: Calculates the electronic energy arising from
            a given configuration. Called by MECI.

DIAT   Utility: Calculates overlap integrals between two atoms in
            general cartesian space. Principal quantum numbers up to 6, and
            angular quantum numbers up to 2 are allowed. Main arguments:
            Atomic numbers and cartesian coordinates in Angstroms of the
            two atoms, Diatomic overlaps (on exit). Called by H1ELEC.

DIAT2   Utility: Calculates reduced overlap integrals between atoms
            of principal quantum numbers 1, 2, and 3, for s and p orbitals.
            Faster than the SS in DIAT. This is a dedicated subroutine, and
            is unable to stand alone without considerable backup. Called

by DIAT.

DIGIT   Utility: Part of READA. DIGIT assembles numbers given a
            character string.

DIHED   Utility: Calculates the dihedral angle between four atoms.
            Used in converting from cartesian to internal coordinates.

DIPOLE   Utility: Evaluates and, if requested, prints dipole components
            and dipole for the molecule or ion. Arguments: Density matrix,
            Charges on every atom, coordinates, dipoles (on exit).
            Called by WRITE and FMAT.

DIPIND   Utility: Similar to DIPOLE, but used by the POLAR calculation
            only.

DOT   Utility: Given two vectors, X and Y, of length N, function DOT
            returns with the dot product X.Y. I.e., if X=Y, then DOT = the
            square of X. Called by FLEPO.

DRC   Main Sequence: The dynamic and intrinsic reaction coordinates
            are calculated by following the mass-weighted trajectories.

DRCOUT   Utility: Sets up DRC and IRC data in quadratic form
            preparatory to being printed.

ENPART   Utility: Partitions the energy of a molecule into its monatomic
            and diatomic components. Called by WRITE when the keyword
            ENPART is specified. No data are changed by this call.

EPSETA   Utility: Calculates the machine precision and dynamic range
            for use by the two diagonalizers.

EXCHNG   Utility: Dedicated procedure for storing 3 parameters and one
            array in a store. Used by SEARCH.

FFHPOL   Utility: Part of the POLAR calculation. Evaluates the
            effect of an electric field on a molecule.

FLEPO   Main Sequence: Optimizes a geometry by minimizing the energy.
            Makes use of the first and estimated second derivatives to
            achieve this end. Arguments: Parameters to be optimized,
            (overwritten on exit with the optimized parameters), Number of
            parameters, final optimized heat of formation. Called by MAIN,
            REACT1, and FORCE.

FMAT   Main sequence: Calculates the exact Hessian matrix for a system
            This is done by either using differences of first derivatives
            (normal mode) or by four full SCF calculations (half electron
            or C.I. mode). Called by FORCE.

FOCK1   Utility: Adds on to Fock matrix the one-center two electron
            terms. Called by ITER only.

FOCK2   Utility: Adds on to Fock matrix the two-center two electron

terms. Called by ITER and DERIV. In ITER the entire Fock matrix
is filled; in DERIV, only diatomic Fock matrices are
constructed.

FOCK2D   Utility: Virtually identical to FOCK2, but with the two-electron
            matrix in double precision. Called by DHC.

FORCE   Main sequence: Performs a force-constant and vibrational
            frequency calculation on a given system. If the starting
            gradients are large, the geometry is optimized to reduce the
            gradient norm, unless LET is specified in the keywords.
            Isotopic substitution is allowed. Thermochemical quantities
            are calculated. Called by MAIN.

FORSAV   Utility: Saves and restores data used in FMAT in FORCE
            calculation. Called by FMAT.

FRAME   Utility: Applies a very rigid constraint on the translations
            and rotations of the system. Used to separate the trivial
            vibrations in a FORCE calculation.

FREQCY   Main sequence: Final stage of a FORCE calculation. Evaluates
            and prints the vibrational frequencies and modes.

GEOUT   Utility: Prints out the current geometry. Can be called at
            any time. Does not change any data.

GETGEO   Utility: Reads in geometry in character mode from specified
            channel, and stores parameters in arrays. Some error-checking
            is done. Called by READ and REACT1.

GETSYM   Utility: Reads in symmetry data. Used by READ.

GMETRY   Utility: Fills the cartesian coordinates array. Data are
            supplied from the array GEO, GEO can be (a) in internal
            coordinates, or (b) in cartesian coordinates. If STEP is
            non-zero, then the coordinates are modified in light of the
            other geometry and STEP. Called by HCORE, DERIV, READ, WRITE,
            MOLDAT, etc.

GOVER   Utility: Calculates the overlap of two Slater orbitals which
            have been expanded into six gaussians. Calculates the
            STP-6G overlap integrals.

GRID   Main Sequence: Calculates a grid of points for a 2-D search
            in coordinate space. Useful when more information is needed
            about a reaction surface.a

H1ELEC   Utility: Given any two atoms in cartesian space, H1ELEC
            calculates the one-electron energies of the off-diagonal
            elements of the atomic orbital matrix.
                    $H(i,j) = -S(i,j)*(beta(i)+beta(j))/2$.
            Called by HCORE and DERIV.

HADDON   Utility: The symmetry operation subroutine, HADDON relates two

geometric variables by making one a dependent function of the other. Called by SYMTRY only.

HCORE   Main sequence: Sets up the energy terms used in calculating the SCF heat of formation. Calculates the one and two electron matrices, and the nuclear energy.
Called by COMPFG.

HELECT   Utility: Given the density matrix, and the one electron and Fock matrices, calculates the electronic energy. No data are changed by a call of HELECT. Called by ITER and DERIV.

HQRII   Utility: Rapid diagonalization routine. Accepts a secular determinant, and produces a set of eigenvectors and eigenvalues. The secular determinant is destroyed.

IJKL   Utility: Fills the large two-electron array over a M.O. basis set. Calls SPCG, and is called by MECI.

INTERP   Utility: Runs the Camp-King converger. q.v.

ITER   Main sequence: Given the one and two electron matrices, ITER calculates the Fock and density matrices, and the electronic energy. Called by COMPFG.

LOCAL   Utility: Given a set of occupied eigenvectors, produces a canonical set of localized bonding orbitals, by a series of 2 x 2 rotations which maximize <psi**4>. Called by WRITE.

LOCMIN   Main sequence: In a gradient minimization, LOCMIN does a line-search to find the gradient norm minimum. Main arguments: current geometry, search direction, step, current gradient norm; on exit: optimized geometry, gradient norm.

MAMULT   Utility: Matrix multiplication. Two matrices, stored as lower half triangular packed arrays, are multiplied together, and the result stored in a third array as the lower half triangular array. Called from PULAY.

MATOUT   Utility: Matrix printer. Prints a square matrix, and a row-vector, usually eigenvectors and eigenvalues. The indices printed depend on the size of the matrix: they can be either over orbitals, atoms, or simply numbers, thus M.O.'s are over orbitals, vibrational modes are over numbers. Called by WRITE, FORCE.

MECI   Main sequence: Main function for Configuration Interaction, MECI constructs the appropriate C.I. matrix, and evaluates the roots, which correspond to the electronic energy of the states of the system. The appropriate root is then returned.
Called by ITER only.

MNDO   Main sequence: MAIN program. MNDO first reads in data using READ, then calls either FLEPO to do geometry optimization,

FORCE to do a FORCE calculation, PATHS for a reaction with a supplied coordinate, NLLSQ for a gradient minimization or REACT1 for locating the transition state. Starts the timer.

MOLDAT   Main Sequence: Sets up all the invariant parameters used during the calculation, e.g. number of electrons, initial atomic orbital populations, number of open shells, etc. Called once by MNDO only.

MOLVAL   Utility: Calculates the contribution from each M.O. to the total valency in the molecule. Empty M.O.'s normally have a negative molecular valency.

MULLIK   Utility: Constructs and prints the Mulliken Population Analysis. Available only for RHF calculations. Called by WRITE.

MULT   Utility: Used by MULLIK only, MULT multiplies two square matrices together.

NLLSQ   Main sequence: Used in the gradient norm minimization.

NUCHAR   Takes a character string and reads all the numbers in it and stores these in an array.

OSINV   Utility: Inverts a square matrix. Called by PULAY only.

PARSAV   Utility: Stores and restores data used in the gradient-norm minimization calculation.

PATHS   Main sequence: Given a reaction coordinate as a row-vector, PATHS performs a FLEPO geometry optimization for each point, the later geometries being initially guessed from a knowledge of the already optimized geometries, and the current step. Called by MNDO only.

PERM   Utility: Permutes n1 electrons of alpha or beta spin among n2 M.O.'s.

POLAR   Utility: Calculates the polarizability volumes for a molecule or ion. Uses 19 SCF calculations, so appears after WRITE has finished. Cannot be used with FORCE, but can be used anywhere else. Called by WRITE

POWSAV   Utility: Calculation store and restart for SIGMA calculation. Called by POWSQ.

POWSQ   Main sequence: The McIver - Komornicki gradient minimization routine. Constructs a full Hessian matrix and proceeds by line-searches Called from MAIN when SIGMA is specified.

PRTDRC   Utility: Prints DRC and IRC results according to instructions. Output can be (a) every point calculated (default), (b) in constant steps in time, space or energy.

PULAY   Utility: A new converger. Uses a powerful
            mathematical non-iterative method for obtaining the SCF Fock
            matrix. Principle is that at SCF the eigenvectors of the Fock
            and density matrices are identical, so [F.P] is a measure of
            the non-self consistency. While very powerful, PULAY is not
            universally applicable. Used by ITER.

QUADR: Utility: Used in printing the IRC - DRC results. Sets up
            a quadratic in time of calculated quantities so that PRTDRC
            can select specific reaction times for printing.

REACT1   Main sequence: Uses reactants and products to find the
            transition state. A hypersphere of N dimensions is centered on
            each moiety, and the radius steadily reduced. The entity of
            lower energy is moved, and when the radius vanishes, the
            transition state is reached. Called by MNDO only.

READ   Main sequence: Almost all the data are read in through READ.
            There is a lot of data-checking in READ, but very little
            calculation. Called by MNDO.

READA   Utility: General purpose character number reader. Used to enter
            numerical data in the control line as " <variable>=n.nnn " where
            <variable> is a mnemonic such as SCFCRT or CHARGE.
            Called by READ, FLEPO, ITER, FORCE, and many other subroutines.

REFS   Utility: Prints the original references for atomic data.
            If an atom does not have a reference, i.e. it has not been
            parameterized, then a warning message will be printed and
            the calculation stopped.

REPP   Utility: Calculates the 22 two-electron reduced repulsion
            integrals, and the 8 electron-nuclear attraction integrals.
            These are in a local coordinate system. Arguments: atomic
            numbers of the two atoms, interatomic distance, and arrays to
            hold the calculated integrals. Called by ROTATE only.

ROTAT   Utility: Rotates analytical two-electron derivatives from
            atomic to molecular frame.

ROTATE   Utility: All the two-electron repulsion integrals, the electron-
            nuclear attraction integrals, and the nuclear-nuclear repulsion
            term between two atoms are calculated here. Typically 100 two-
            electron integrals are evaluated.

RSP   Utility: Rapid diagonalization routine. Accepts a secular
            determinant, and produces a set of eigenvectors and
            eigenvalues. The secular determinant is destroyed.

SCHMIB   Utility: Part of Camp-King converger.

SCHMIT   Utility: Part of Camp-King converger.

SEARCH   Utility: Part of the SIGMA and NLLSQ gradient minimizations.
            The line-search subroutine, SEARCH locates the gradient

minimum and calculates the second derivative of the energy
in the search direction. Called by POWSQ and NLLSQ.

SECOND   Utility: Contains VAX specific code. Function SECOND
            returns the number of CPU seconds elapsed since an arbitrary
            starting time. If the SHUTDOWN command has been issued,
            the CPU time is in error by exactly 1,000,000 seconds, and
            the job usually terminates with the message "time exceeded".

SET   Utility: Called by DIAT2, evaluates some terms used in overlap
            calculation.

SETUPG   Utility: Sets up the Gaussian expansion of Slater orbitals
            using a STO-6G basis set.

SOLROT   Utility: For Cluster systems, adds all the two-electron
            integrals of the same type, between different unit cells, and
            stores them in a single array. Has no effect on molecules.

SPCG   Utility: Calculates two-electron integral between any four M.O.'s
            at the MNDO or MINDO/3 level. Called by MECI and WRITE.

SPLINE   Utility: Part of Camp-King converger.

SS   Utility: An almost general Slater orbital overlap calculation.
            Called by DIAT.

SWAP   Utility: Used with FILL=, SWAP ensures that a specified
            M.O. is filled. Called by ITER only.

SYMTRY   Utility: Calculates values for geometric parameters from known
            geometric parameters and symmetry data. Called whenever GMETRY
            is called.

THERMO   Main sequence: After the vibrational frequencies have been
            calculated, THERMO calculates thermodynamic quantities such as
            internal energy, heat capacity, entropy, etc, for translational,
            vibrational, and rotational, degrees of freedom.

TIMBGN   VAX-specific code for determining CPU time.

TQL2   Utility: Part of the RSP.

TQLRAT   Utility: Part of the RSP.

TRBAK3   Utility: Part of the RSP.

TRED3   Utility: Part of the RSP.

UPDATE   Utility: Given a set of new parameters, stores these
            in their appropriate arrays. Invoked by EXTERNAL.

VECPRT   Utility: Prints out a packed, lower-half triangular matrix.
            The labeling of the sides of the matrix depend on the matrix's

size: if it is equal to the number of orbitals, atoms, or other.
Arguments: The matrix to be printed, size of matrix. No data
are changed by a call of VECPRT.

WORD   Utility: Part of WRTKEY, checks keywords for recognition.
          If the keyword is recognised, it is ignored. Any words
          not recognised will be flagged and the job stopped.

WRITE   Main sequence: Most of the results are printed here. All
          relevant arrays are assumed to be filled. A call of WRITE only
          changes the number of SCF calls made, this is reset to zero.
          No other data are changed. Called by MAIN, FLEPO, FORCE.

WRTKEY   Main Sequence: Prints all keywords and checks for
          compatability and to see if any are not recognised.
          WRTKEY can stop the job if any errors are found.

XYZINT   Utility: Converts from cartesian coordinates into internal.

XYZGEO   XYZINT sets up its own numbering system, so no connectivity
          is needed.

APPENDIX D

HEATS OF FORMATION OF SOME MNDO, PM3 AND AM1 COMPOUNDS

In order to verify that MOPAC is working correctly, a large number of tests need to be done. These take about 45 minutes on a VAX 11-780, and even then many potential bugs remain undetected. It is obviously impractical to ask users to test MOPAC. However, users must be able to verify the basic working of MOPAC, and to do this the following tests for the elements have been provided.

Each element can be tested by making up a data-file using estimated geometries and running that file using MOPAC. The optimized geometries should give rise to heats of formation as shown. Any difference greater than 0.1 Kcal/mole indicates a serious error in the program.

Caveats

1. Geometry definitions must be correct.

2. Heats of formation may be too high for certain compounds. This is due to a poor starting geometry trapping the system in an excited state. (Affects ICl at times)

Element	Test Compound	Heat of Formation		
		MNDO	AM1	PM3
Hydrogen	CH_4	-11.9	-8.8	-13.0
Lithium	LiH	+23.2		
Beryllium	BeO	+38.6		
Boron	BF_3	-261.0		
Carbon	CH_4	-11.9	-8.8	-13.0
Nitrogen	NH_3	-6.0	-7.3	-3.1
Oxygen	CO_2	-75.0	-79.8	-85.0
Fluorine	CF_4	-214.2	-225.7	-225.1
Aluminium	AlF	-83.6		-50.1
Silicon	SiH	+90.2		+94.6
Phosphorus	PF_3	-229.3		-252.2
Sulfur	H_2S	+1.7		-0.9
Chlorine	HCl	-15.3	-24.6	-20.5
Germanium	GeF	-16.4		
Bromine	HBr	+3.6	-10.5	+5.3
Tin	SnF	-20.4		
Iodine	ICl	-6.7	-4.6	+10.8
Mercury	HgO	+101.6		
Lead	PbF	-22.6		

APPENDIX E

REFERENCES

On MNDO

"Ground States of Molecules. 38. The MNDO Method. Approximations and Parameters.", M.J.S. Dewar, W.Thiel, J. Am. Chem. Soc., 99, 4899, (1977).

Original References for Elements Parameterized in MNDO

H   M.J.S. Dewar, W. Thiel, J. Am. Chem. Soc., 99, 4907, (1977).

Li   Parameters taken from the MNDOC program, written by Walter Thiel, Quant. Chem. Prog. Exch. No. 438; 2, 63, (1982)

Be   M.J.S. Dewar, H.S. Rzepa, J. Am. Chem. Soc, 100, 777, (1978)

B   M.J.S. Dewar, M.L. McKee, J. Am. Chem. Soc., 99, 5231, (1977).

C   M.J.S. Dewar, W. Thiel, J. Am. Chem. Soc., 99, 4907, (1977).

N   M.J.S. Dewar, W. Thiel, J. Am. Chem. Soc., 99, 4907, (1977).

O   M.J.S. Dewar, W. Thiel, J. Am. Chem. Soc., 99, 4907, (1977).

F   M.J.S. Dewar, H.S. Rzepa, J. Am. Chem. Soc., 100, 58, (1978).

Al   L.P. Davis, R.M. Guidry, J.R. Williams, M.J.S. Dewar, H.S. Rzepa J. Comp. Chem., 2 433, (1981).

Si   (a) M.J.S. Dewar, M.L. McKee, H.S. Rzepa, J. Am. Chem. Soc., 100, 3607 (1978). *
     (c) M.J.S. Dewar, J. Friedheim, G. Grady, E.F. Healy, J.J.P. Stewart, Organometallics, 5, 375 (1986).

P   M.J.S. Dewar, M.L. McKee, H.S. Rzepa, J. Am. Chem. Soc., 100, 3607 (1978).

S   (a) M.J.S. Dewar, M.L. McKee, H.S. Rzepa, J. Am. Chem. Soc., 100, 3607 (1978). *
     (b) M.J.S. Dewar, C. H. Reynolds, J. Comp. Chem., 7, 140 (1986).

Cl   (a) M.J.S. Dewar, M.L. McKee, H.S. Rzepa, J. Am. Chem. Soc.,
      100, 3607 (1978). *
      (b) M.J.S. Dewar, H.S. Rzepa, J. Comp. Chem., 4, 158, (1983)

Ge   M.J.S. Dewar, G.L. Grady, E.F. Healy, Organometallics, 6,
      186 (1987).

Br   M.J.S. Dewar, E.F. Healy, J. Comp. Chem., 4, 542, (1983)

I   M.J.S. Dewar, E.F. Healy, J.J.P. Stewart, J. Comp. Chem.,
      5, 358, (1984)

Sn   M.J.S. Dewar, G.L. Grady, J.J.P. Stewart, J. Am. Chem. Soc.,
      106, 6771 (1984).

Hg   M.J.S. Dewar, G.L. Grady, K. Merz, J.J.P. Stewart,
      Organometallics, 4, 1964, (1985).

Pb   M.J.S. Dewar, M. Holloway, G.L. Grady, J.J.P. Stewart,
      Organometallics, 4, 1973, (1985).

  * - Parameters defined here are obsolete.

On MINDO/3
            Part XXVI, Bingham, R.C., Dewar, M.J.S., Lo, D.H,
            J. Am. Chem. Soc., 97, (1975).

On AM1
            "AM1: A New General Purpose Quantum Mechanical Molecular
            Model", M.J.S. Dewar, E.G. Zoebisch, E.F. Healy,
            J.J.P. Stewart, J. Am. Chem. Soc., 107, 3902-3909 (1985).

On PM3
            "Optimization of Parameters for Semi-Empirical Methods I-Method",
J.J.P. Stewart, J. Comp. Chem. (In Press, expected date of publication,
February 1989).
            "Optimization of Parameters for Semi-Empirical Methods
II-Applications, J.J.P. Stewart, J. Comp. Chem. (In Press, expected date of
publication, February 1989).

            Original References for Elements Parameterized in AM1
H   M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am.
      Chem. Soc., 107, 3902-3909 (1985).
B   M.J.S. Dewar, C Jie, E. G. Zoebisch, Organometallics,
      7 513-521 (1988).
C   M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am.
      Chem. Soc., 107, 3902-3909 (1985).
N   M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am.
      Chem. Soc., 107, 3902-3909 (1985).
O   M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am.
      Chem. Soc., 107, 3902-3909 (1985).
F   No reference available at this time.
Cl   No reference available at this time.
Br   No reference available at this time.

I   No reference available at this time.
 (see also PARASOK for the use of MNDO parameters for other elements)

On Shift
        "The Dynamic 'Level Shift' Method for Improving the
        Convergence of the SCF Procedure", A. V. Mitin, J. Comp.
        Chem. 9, 107-110 (1988).

On Half-Electron
        "Ground States of Conjugated Molecules.
        IX. Hydrocarbon Radicals and Radical Ions", M.J.S. Dewar,
        J.A. Hashmall, C.G. Venier, J.A.C.S. 90, 1953 (1968).
        "Triplet States of Aromatic Hydrocarbons", M.J.S. Dewar,
        N. Trinajstic, Chem. Comm., 646, (1970).
        "Semiempirical SCF-MO Treatment of Excited States of
        Aromatic Compounds" M.J.S. Dewar, N. Trinajstic,
        J. Chem. Soc., (A), 1220, (1971).

On Pulay's Converger
        "Convergence Acceleration of Iterative Sequences.
         The Case of SCF Iteration", Pulay, P., Chem. Phys. Lett.,
        73, 393, (1980).

On Pseudodiagonalization
        "Fast Semiempirical Calculations",
        Stewart. J.J.P., Csaszar, P., Pulay, P., J. Comp. Chem.,
        3, 227, (1982).

On Localization
        "A New Rapid Method for Orbital Localization."
        P.G. Perkins and J.J.P. Stewart, J.C.S. Faraday
        (II) 77, 000, (1981).

On Diagonalization
        Beppu, Y., Computers and Chemistry,
        Vol.6 (1982).

On MECI
        "Molecular Orbital Theory for the Excited States of
         Transition Metal Complexes", D.R. Armstrong, R. Fortune,
         P.G. Perkins, and J.J.P. Stewart, J. Chem. Soc., Faraday
         2, 68 1839-1846 (1972)

On Broyden-Fletcher-Goldfarb-Shanno Method
     Broyden, C. G., Journal of the Institute for Mathematics and
                    Applications, Vol. 6 pp 222-231, 1970.
     Fletcher, R., Computer Journal, Vol. 13, pp 317-322, 1970.

     Goldfarb, D. Mathematics of Computation, Vol. 24, pp 23-26, 1970.

     Shanno, D. F. Mathematics of Computation, Vol. 24, pp 647-656
                    1970.

  See also summary in

Shanno, D. F., J. of Optimization Theory and Applications
        Vol.46, No 1 pp 87-94 1985.

On Polarizability

        "A New Procedure for Calculating Molecular
        Polarizabilities: Applications Using MNDO."
        M.J.S. Dewar, J.J.P. Stewart, Chem. Phys.
        Lett. 111 416 (1984).

On Thermodynamics
        "Ground States of Molecules. 44 MINDO/3 Calculations of
        Absolute Heat Capacities and Entropies of Molecules
        without Internal Rotations." Dewar, M.J.S., Ford, G.P.,
        J. Am. Chem. Soc., 99, 7822 (1977).

On SIGMA Method
        Komornicki, A., McIver, J. W., Chem. Phys. Lett., 10,
        303, (1971).
        Komornicki, A., McIver, J. W., J. Am. Chem. Soc., H
        94, 2625 (1971)

On Molecular Orbital Valency
        "Valency and Molecular Structure", Gopinathan, M. S.,
        Siddarth, P., Ravimohan, C., Theor. Chim. Acta
        70, 303 (1986).

On Bonds
        "Bond Indices and Valency", Armstrong, D.R.,
        Perkins, P.G., Stewart, J.J.P., J. Chem. Soc.,
        Dalton, 838 (1973).
For a second, equivalent, description, see also Gopinathan, M. S., and
Jug, K., Theor. Chim. Acta, 63, 497 (1983).

On Locating Transition States
        "Location of Transition States in Reaction Mechanisms",
        M.J.S. Dewar, E.F. Healy, J.J.P. Stewart, J. Chem. Soc.,
        Faraday Trans. 2, 3, 227, (1984)

On Dipole Moments for Ions
        "Molecular Quadrupole Moments", A.D. Buckingham, Quarterly
        Reviews, 182 (1958 or 1959)

INDEX